

Amendments to the Drawings:

The attached sheet of drawings includes changes to Figs. 1, 3, and 4. This sheet, which includes Figs. 1-16, replaces the original sheet including Figs. 1-16. In Figs. 1, 3, and 4, previously omitted SEQ ID NO:s have been added.

Attachment: Replacement Sheet

Annotated Sheet Showing Changes

REMARKS/ARGUMENTS


In the specification, the paragraph on page 27, beginning with line 20, has been amended to add the previously omitted SEQ ID NOs to each nucleotide sequence that is more than ten nucleotides in length. Appendices I and II have been replaced with new Appendices I and II, which merely add the previously omitted SEQ ID NOs to each nucleotide sequence that is more than ten nucleotides in length.

In amended Figure 1, the previously omitted SEQ ID NOs have been added to each nucleotide sequence more than ten nucleotides in length and to each amino acid sequences more than four amino acids in length. Similarly, in amended Figures 3 and 4, the previously omitted SEQ ID NOs have been added to each nucleotide sequence more than ten nucleotides in length.

Applicants respectfully request entrance of the above amendments. In view of the above remarks, early notification of a favorable consideration and allowance of all claims is respectfully requested.

If after this amendment there are issues remaining which discussion could advance prosecution, Applicants respectfully request that the examiner call the undersigned attorney at the phone number listed.

Dated: February 2, 2004

Respectfully submitted,
By: 

Laurence H. Posorske
Registration No. 34,698
David A. Kelly
Registration No. 53,106

HUNTON & WILLIAMS
1900 K Street, N.W.
Suite 1200
Washington, D.C. 20006-1109
Telephone: (202) 955-1500
Facsimile: (202) 778-2201

Attachments



Annotated Sheet Showing Changes

Figure 1
~~FIGURE 1~~

ID-65

Clone 3-60

5 (SEQ ID No: 1)
GTGTTTATGATGAAAAAAGGACAAGTAAATGATACTAAGCAA
TCTTACTCTCTACGTAAATATAAATTTGGTTTAGCATCAGTAA
TTTTAGGGTCATTCATAATGGTCACAAGTCCTGTTTTTGCGGA
TCAAACATACATCGGTTCAAGTTAATAATCAGACAGGCACTAG
10 TGTGGATGCTAATAATTCTTCCAATGAGACAAGTGCGTCAAGT
GTGATTACTTCCAATAATGATAGTGTTCAAGCGTCTGATAAAG
TTGTAAATAGTCAAAATACGGCAACAAAGGACATTACTACTC
CTTTAGTAGAGACAAAGCCAATGGTGGAAAAAACATTACCTG
AACAAGGGAATTATGTTTATAGCAAAGAAACCGAGGTGAAAA
15 ATACACCTTCAAAATCAGCCCCAGTAGCTTTCTATGCAAAGA
AAGGTGATAAAGTTTTCTATGACCAAGTATTTAATAAAGATA
ATGTGAAATGGATTTCATATAAGTCTTTTGGTGGCGTACGTCG
ATACGCAGCTATTGAGTCACTAGATCCATCAGGAGGTTTCAGA
GACTAAAGCACCTACTCCTGTAACAAATTCAGGAAGCAATAA
20 TCAAGAGAAAATAGCAACGCAAGGAAATTATACATTTTCACA
TAAAGTAGAAGTAAAAAATGAAGCTAAGGTAGCGAGTCCAA
CTCAATTTACATTGGACAAAGGAGACAGAATTTTTTTACGACC
AAATACTAACTATTGAAGGAAATCAGTGGTTATCTTATAAATC
ATTCAATGGTGTTCGTCGTTTTGTTTTGCTAGGTAAAGCATCTT
25 CAGTAGAAAAAACTGAAGATAAAGAAAAAAGTGTCTCCTCAAC
CACAAGCCCGTATTACTAAAACCTGGTAGACTGACTATTTCTAA
CGAAACAACACTACAGGTTTTGATATTTTAATTACGAATATTA
GATGATAACGGTATCGCTGCTGTTAAGGTACCGGTTTGGACTG
AACAAGGAGGGCAAGATGATATTAATGGTATACAGCTGTAA
30 CTACTGGGGATGGCAACTACAAAGTAGCTGTATCATTTGCTG
ACCATAAGAATGAGAAGGGTCTTTATAATATTCATTTATACTA
CCAAGAAGCTAGTGGGACACTTGTAGGTGTAACAGGAACTAA
AGTGACAGTAGCTGGAACCTAATTCTTCTCAAGAACCTATTGA
AAATGGTTTACCAAAGACTGGTGTTTATAATATTATCGGAAGT
35 ACTGAAGTAAAAAATGAAGCTAAAATATCAAGTCAGACCCAA
TTTACTTTAGAAAAAAGGTGACAAAATAAATTATGATCAAGTA
TTGACAGCAGATGGTTACCAGTGGATTCTTACAAATCTTATA
GTGGTGTTCGTCGCTATATTCCTGTGAAAAAGCTAACTACAAG
TAGTGAAAAAGCGAAAGATGAGGCGACTAAACCGACTAGTTA
40 TCCCAACTTACCTAAAACAGGTACCTATACATTTACTAAAAC
GTAGATGTGAAAAGTCAACCTAAAGTATCAAGTCCAGTGGAA
TTTAATTTTCAAAAGGGTGAAAAAATACATTATGATCAAGTGT
TAGTAGTAGATGGTCATCAGTGGATTTCATACAAGAGTTATTC
CGGTATTCGTCGCTATATTGAAATTTAA

(SEQ ID NO: 2)

MFMMKKKGQVNDTKQSYSLRKYKFGLASVILGSFIMVTSPVFADQTTSVQVNN
QTGTSVDANNSSNETSASSVITSNND SVQASDKVVNSQNTATKDITTPLVETK
5 PMVEKTLPEQGN YVYSKETEVKNTPSKSAPVAFYAKKGDKVFYDQVFNKDN
VKWISYKSFGGVRRYAAIESLDPSGGSETKAPTPVTNSGSNNQEKIATQGN Y
FSHKVEVKNEAKVASPTQFTLDKGDRIFYDQILTIEGNQWLSYKSFNGVRRFV
LLGKASSVEKTEDKEKVSPQPQARITKTGRLTISNETTTGFDILITNIKDDNGIA
AVKVPVWTEQGGQDDIKWYTA VTTGDGNYKVAVSFADHKNEKGLYNIHLY
10 YQEASGTLVGVTGKVTVAGTNSSQEPIENGLPKTGVYNIIGSTEVKNEAKISS
QTQFTLEKGD KINYDQVLTADGYQWISYKSYSGVRRYIPVKKLTTSSEKAKDE
ATKPTSYPNLPKTGT YTFKTVDVKSQPKVSSPVEFNFQKGEKIH YDQVLVVD
GHQWISYKSYSGIRRYIEI*

15

Sequence description

A) Length: 1642 bp - 547 aa (full length gene)

B) Sequence Characteristics:

Potential leader peptide sequence

20

Orf is preceded by a potential Shine-
Dalgarno sequence.

ID-66

25

Clone 3-5

(SEQ ID NO: 3)

ATGATATTGAGACGTCGA ACTATTGTTTTATGGCAACTGGGTATCGCCATT
TCTCTCATTCTTAGTATTCTAGCCTTAAATCTTTATTTCCATAGTACTCCCTT
30 GCAAACCAATGCAGCTTTACGGAACCTTGCTCCTTCATTAAACCATCTTTTT
GGGACAGATGGTTTAGGTAGGGATATGTTTGT CAGAACGATTAAAGGACT
TTATTTCTCTCTACAAGTCGGCTTATTAGGTGCCCTTATGGGGGGTCATTCTG
GCGACAGTTTTTGGAGTGCTTGCAGGTTTAGGAAATAGCATTATTGATAAA
ATAATAGCATGGTTAGTTGATTTGTTTATTGGTATGCCTCATTTGATTTTTA
35 TGATTCTCATTTCTTTTGTTGTTGGGAAAGGTGCTCAAGGGGGTCATCATTGC
AACGGCTGTTACACATTGGCCTTCTTTAGCAAGGCTTATCCGCAATGAAGT
CTATCATCTAAAGAATAAAGAATTTGTCCA ACTTTCTAAAAGTATGGGAAA
AACGCCTTATTATATTGTGAGGCATCATATCCTGCCTTTGATTGCTTCTCAA
ATTTTCATTGGTTTTATCCTCTTATTTCCACATGTCATCCTACATGAAGCAT
40 CAATGACTTTCTTAGGATTTGGGCTCTCTGCCGAACAACCTTCGGTTGGTA
TCATTCTGTCAGAGGCAGCTAAGCATATCTCTCTTGGA AATTGGTGGTTGG
TTATCTTTCCAGGACTTTATCTTATTTTGGTTGTCAATGCATTTGATACTAT
CGGAGAATCTTTAAAGAACTCTTTTACCCTCAA ACTGATCATTTTTAG

(SEQ ID NO: 4)

MILRRRTIVLWQLGIAISLILSILALNLYFHSTPLQTNAALRNLAAPSLNHLFGTD
GLGRDMFVRTIKGLYFSLQVGLLGALMGVILATVFGVLAGLGNSIIDKIIAWL
VDLFIGMPHLIFMILISFVVGKGAQGVIIATAVTHWPSLARLIRNEVYHLKNKE
5 FVQLSKSMGKTPYYIVRHHILPLIASQIFIGFILLFPHVILHEASMTFLGFGLSAE
QPSVGIILSEAAKHISLGNWWLVIFPGLYLILVVNAFDTIGESLKKLFYPQTDHF
*

10 Sequence description

A) Length: 822 bp - 274 aa (full length gene)

B) Sequence Characteristics:

Potential leader peptide sequence

15 Orf is preceded by a potential Shine-
Dalgarno sequence.

ID-78

20

Clone 3-5b

(SEQ ID NO: 5)

ATGACAGAAACATTATTAAGCATTAAAGACCTCTCCATCACCTTCACTCAA
TACGGAAGATTTTAAACCATTTCATCAACACCGATACAAGCGCTGA
25 ATTTAGAAATTAAAAAAGGTGAGTTATTAGCTATTATAGGTGCTAGTGGTT
CGGGGAAGAGTTTATTAGCACATGCTATTATGGATATTCTTCCTAAAAATG
CATCTGTAACAGGAGATATGATTTATCGTGGTCAATCACTAAATTCTAAAC
GCATTAAACAGTTGCGAGGAAAAGATATTACGTTGATTCCACAATCAGTTA
ATTATTTAGATCCATCTATGAAAGTCAAACATCAGGTGCGCTTAGGTATCT
30 CAGAAAATTCAAAGGCTACTCAAGAAGGATTGTTTCAACAGTTTGGTTTAA
AAGAAAGTGATGGTGACTTGGATCCTTTCCAACCTTTCTGGCGGAATGCTCC
GACGTGTTTTGTTTACAACGTGTATTAGTGATAAGGTTTCTTTGATTATTGC
GGATGAGCCCACCCCTGGATTACATCCAGATGCTCTGCAAATGGTTTTAGA
CCAACACGCTCCTTTGCAGATAAAGGAATAAGCGTTATATTTATCACTCA
35 TGATATTGTAGCAGCTAGTCAAATTGCTGATCGTATTACTATTTTAAAGA
GGGAAAAGCTATTGAAACAGCTCCAGCTAGTTTCTTTAGCGGAAATGGAG
AGCAGTTACAAACAGAATTTGCTAGAAGTTTATGGCGCTCTCTCCACAGC
AAGAATTTTGAAGGAGTTACTCATGACCTTAGAGGCTAA

(SEQ ID NO: 6)

40 MTETLLSIKDLSTFTQYGRFLKPFQSTPIQALNLEIKKGELLAIIIGASGSGKSL
AHAIMDILPKNASVTGDMIYRGQSLNSKRIKQLRGKDITLIPQSVNYLDPSMK
VKHQVRLGISENSKATQEGLFQQFGLKESDGDLDPFQLSGGMLRRVLFTTCIS
DKVSLIIADEPTPGLHPDALQMVLQDLRSFADKGISVIFITHDIVAASQIADRITI
FKEGKAIETAPASFFSGNGEQLQTEFARSLWRSLPQQEFLKGVTHDLRG*

5

10

15

20

25

35

40

This gene was not isolated using the LEEP system. However in determining a full length gene sequence for ID-76, this gene was identified upstream.

ID-80

5

Clone 2-17

(SEQ ID NO: 9)

10 TTGCGGACAATTACGTTCAAACACAATGAAACGCGATCGTCAAAAAGCGA
AGGTAGGGCGGTAATGCTTAAAAGATTATTTACTGAAGATGGGGAATTGA
CAAAGATTAGTCGTCGTTTCGTTTGGATGTTAGTGGTTATCTATTGTCTTAT
TATTGTCAGGATGTGTTTTGGGCCTCAAATTATGATTGAGGGGGTATCAAC
TCCGAATGTTTCAGCGCTTCGGAAGAATTGTAGCTCTTTTAGTACCATTTAA
TTCTTTTCGTAGTTTAGATCAGCTAACTAGCTTTAAAGAGATTTTTTGGGTT
ATTGGTCAAAATGTAGTGAATATTTTACTGCTGTTTCCTCTCATTATAGGGT
15 TACTATCCCTAAAGCCAAGTTTACGGAAATATAAAAGCGTTATATTACTTG
CTTCTTGATGTCTCTTTTCATAGAGTGTACTCAAGTTGTTTTAGATATTTT
AATAGATGCTAATCGGGTTTTTGAAATCGACGATCTATGGACAAATACCTT
AGGCGGTCCTTTCGCCCTATGGAGTTATCGAAACATAAAAGGTTGGCTTCT
AACTATTAGAAAATGA

20

(SEQ ID NO: 10)

MRTITFKHNETRSSKSEGRAVMLKRLFTEDGELTKISRRFVWMLVVIYCLIIVR
MCFGPPQIMIEGVSTPNVQRFGRIVALLVPFNSFRSLDQLTSFKEIFWVIGQNVV
NILLFPLIIGLLSLKPSLRKYKSVILLAFLMSLFIECTQVVLDILIDANRVFEIDD
LWTNTLGGPALWSYRNIKGWLLTIRK*

25

Sequence description

30 A) Length: 579 bp - 193 aa (full length gene)
B) Sequence Characteristics:
Possesses a potential leader peptide sequence
No obvious Shine-Dalgarno, but the 'TTG' codon
may not be the actual translation start point.
A methionine (ATG) that occurs ~22 codons
35 downstream of the 'TTG' is preceded by a
potential Shine-Dalgarno sequence and may
represent the actual start codon.

ID 81

40

Clone 3-1

(SEQ ID NO: 11)

TTGAAAAATTTAAATCGTTATGTAGTTGCGGTTTCTGGAGTCGTTTTACATT
 TAATGCTAGGATCAACTTATGCTTGGAGTGTGTTTCGTAACCCAATTATCT
 CAGAGACTGGTTGGGATATTTTCATCAGTTTCATTCGCTTTTAGTTTGGCTAT
 TTTTGTCTAGGAATGTCTGCAGCTTTTATGGGACACTTAGTAGAGCGTTTT
 5 GGTCCTAGGATAATGGGAATGATTTCTGCTATTTTATATGGAGCAGGGAAT
 GTGTTAACAGGCTTAGCCATTGAACTCAGCAGTTATGGTTACTGTATGTT
 GCATACGGTATTTTAGGAGGAATCGGACTTGGTTCAGGTTATATTACTCCA
 GTATCGACTATTATTAAATGGTTTCCTGATAGGAGGGGACTAGCAACAGG
 ATTCGCTATTATGGGATTTGGCTTTGCTTCTTTAGTAACAAGTCCGCTTGCA
 10 CAATCCTTACTGATTAGGATTGGTGTGGGTAAAACGTTTTATATTTTGGGA
 TTAGTATATTTTTTTGTCATGATGATTGCCTCACAATTTATTAAACAACCAC
 CTCAGGAAAAAATAACTATTTTGACTCACGATGGTAAAAAGAATGCTATG
 AATTCACAAATTATCACTGGATTAAAAGCAAACGTCGCTATAAAATCAAA
 AACCTTTTACATCATTTGGTTGACCTTGTTTATTAATATTTTCGTGTGGCTTA
 15 GGTTTAATATCAGCAGCTTCACCAATGGCACAAGATTTAGCAGGCTATTCC
 GCAGAATCTGCAGCCTTATTAGTAGGGGTACTAGGGATATTTAACGGTTTT
 GGACGTCTGTTATGGGCAAGTCTCTCTGACTACATTGGACGCCCGTTGACC
 TTTATAATATTATTTATTGTGAACCTTATTATGACTTCTAGTTTATTTTTGTC
 ATTCAATGCTATTGTATTTGCAATAGCGATGTCTATTTTAATGACTTGTTAT
 20 GGTGCAGGTTTTTCCTTATTACCTGCTTATCTAAGTGATATTTTTTGGAAACAA
 AGGAATTAGCTACTTTACATGGTTATAGTTTAAACAGCATGGGCAATAGCAG
 GTCTGTTTGGGCCCCTATTGTTATCAAAGACATATTCATGGGGAAATTCCT
 ATCAATTGACATTAATGGTTTTTGGTTTTTTTATTCTTATTCGGATTATTGTTA
 TCTCTATATTTAAGAAAATTAACAACCTAAAGTTGTGTAG
 25 (SEQ ID NO: 12)
 LKNLNRVYVAVSGVVLHMLGSTYAWSVFRNPIISETGWDISSVSFAFSLAIFC
 LGMSAAFMGHLVERFGPRIMG MISAILYGAGNVLTGLAIETQQLWLLYVAYG
 ILGGIGLGSYITPVSTIIKWFPDRRGLATGFAIMGFGFASLVTSP LAQSLLIRIG
 VGKTFYILGLVYFFVMMIASQFIKQPPQEKITILTHDGKKNAMNSQIITGLKAN
 30 VAIKSKTFYIIWLTLFINISCGLGLISAASPM AQDLAGYSAESAALLVGVLGIFN
 GFGRLLWASLSDYIGRPLTFILFIVNFIMTSSLFSLFNAIVFAIAMSILMTCYGA
 GFSLLPAYLSDIFGTKELATLHGYSLTAWAIAGLFGPLLLSKTYSWGNSYQLTL
 MVFGFLFLFGLLLSLYLRKLTTKVV*

35

Sequence description:

- 40 A] Length 1221 bp - 407 a.a (full length gene).
 B] TTG start codon with Shine-Dalgarno sequence upstream. Obvious signal peptide, with hydropathy plot exhibiting many possible membrane spanning regions, indicating protein to be transmembrane.

5 ID-82

Clone 48

10 (SEQ ID NO: 13)
ATGGCAGATAAAAAACAGAACATTTAAACTTGTAGGTGCAGGATCTTCTAG
CACACAAGAAAAAATTGAAAAGCCTGCTCTTTCGTTTATGCAAGATGCGTG
GCGTCGCTTGAAAAAAAACAAATTAGCAGTAGTTTCACTCTATTTATTAGC
TCTTTTACTTACTTTTTCGTTAGCCTCAAATTTATTTGTAACCTCAGAAGGAT
GCTAATGGGTTTGATTTCGAAAAAAGTAACGACATATCGCAACTTACCACCT
15 AAATTGAGTTCAAACCTTCCTTTTTTGAATGGTAGCATTAATCCATCA

(SEQ ID NO: 14)
MADKNRTFKLVGAGSSSTQEKIEKPALSFMQDAWRRLKKNKLAVVSLYLLA
LLLTFSLASNLFVTQKDANGFDSKKVTTYRNLPPKLSSNLPFWNGSINPS

20

Sequence description:

25 A] Current length is 303 bp - 101 aa
B] No obvious signal peptide but Shine
Dalgarno sequence upstream of the ATG start
codon. Not identified directly using the LEEP system but was found
directly downstream of ID-34 described in WO 00/06736.

—————>

30 ID-83

Clone 98

35 (SEQ ID NO: 15)
ATGAAAATAGTAGTACCAGTAATGCCTCGCAGTCTTGAAGAGGCTCAAGA
AATAGATTTATCAAAATTTGATAGTGTTGATATTATTGAATGGCGAGCTGA
TGCCTTACCAAAGGATGACATTATTAATGTAGCTCCAGCTATTTTTTGAGAA
ATTCGCAGGTCATGAAATTATTTTTACTTTTCGTACAACGCGTGAAGGTGG
40 TAATATTGTCTTATCTGATGCTGAGTATGTTGAGTTAATCCAGAAAATTAA
TTCTATCTACAATCCAGATTATATTGATTTTGAGTATTTTTTCACATAAAGAA
GTTTTTCAAGAAATGCTAGAATTTCCAAATTTAGTCCTGTCTTATCACAATT
TTCAAGAGACACCGGAGAATATTATGGAGATATTTTCAGAATTAACAGCC
CTAGCACCACGAGTTGTGAAAATCGCAGTAATGCCAAAGAATGAACAAGA

5 TGTCTTAGACGTTATGAATTACACTCGCGGTTTCAAGACTATTAATCCTGA
TCAAGTTTATGCGACGGTATCTATGAGTAAAATTGGACGTATTTCTCGTTTT
GCTGGTGATGTAACCTGGATCTAGTTGGACATTTGCATATTTAGATTCATCT
ATCGCACCCGGACAAATTACTATTTTCAGAGATGAAGCGTGTCAAAGCATT
GCTTGACGCTGACTGA

(SEQ ID NO: 16)
10 MKIVVPVMPRSLLEEAQEIDLSKFDSVDIIEWRADALPKDDIINVAPAIFEFK
HEIIFTFRTTREGGNIVLSDAEYVELIQKINSIYNPDYIDFEYFSHKEVFQEMLEF
PNLVLSYHNFQETPENIMEIFSELTALAPRVVKIAVMPKNEQDVLDVMNYTRG
FKTINPDQVYATVSMSKIGRISRFAGDVTGSSWTFAYLDSSIAPGQITISEMKRV
KALLDAD*

15 Sequence description:

- A] Length 678 bp, 225 aa (full length gene)
- B] No obvious signal peptide, but there is a
Shine Dalgarno immediately upstream of ORF.

20 ID-84

25 Clone RS-52

(SEQ ID NO: 17)
30 ATGAAAGACTTATTTGCAACAACAGAAGCATCATCAAGGAAACAGGAACA
AGATAGAATTGTCAATTACATAAAACAACATGTTGAGTTAACAAATGGTA
ATCAAATAAAAAAAAAATTGAGTTTATCGACTTTCAAAAAAAAAATGAGATGACA
GGTACATGGGGAATTTCTACTAAAATTAATGAACAATTTTCGATTAGTTTT
TCTGAAGATAGAATTGGTGGTAAACTTAGAGCATTAGGATATCAACCGAA
TGAAATAGGTTTTTCAAAGGACATCAATAGTAATAATCAAAATGTTAATGA
TATTGAAGTGATTTATATGAAGAAAGAATAG

(SEQ ID NO: 18)
35 MKDLFATTEASSRKQEQRIVNYIKQHVELTNGNQIKKIEFIDFQKNEMTGTW
GISTKINEQFSISFSEDRIGGKLRALGYQPNEIGFSKDINSNNQNVNDIEVIYMK
KE*

40 Sequence description:

- A] length: 333 bp - 111 aa (partial sequence)
- B] No obvious Shine Dalgarno sequence upstream
of the ATG start codon, and no obvious signal
peptide within the protein.

ID-85

5

Clone RS-53

(SEQ ID NO: 19)
10 ATGAAAAAACGTATATGGTATTTGATAATAATAATCACAGTAATTTTAGGA
GGACTAGCCATGAAAAACTTATTTGCAACAACAGAAGCATCATCAAGGAA
ACAGGAACAAGATAGAATTGTCAATTACATAAAACAACATGTTGAGTTAA
CAAATGGTAATCAAATAAAAAAAATTGAGTTTATCGACTTTCAAAAAAAT
GAGATGACAGGTACATGGGGAATTTCTACTAAAATTAATGAACAATTTTCG
15 ATTAGTTTTTCTGAAGATAGAATTGGTGGTAAACTTAGAGCATTAGGATAT
CAACCGAATGAAATAGGTTTTTCAAAGGACATCAATAGTAATAATCA
(SEQ ID NO: 20)
MKKRIWYLIITVILGGLAMKNLFATTEASSRKQEQDRIVNYIKQHVELTNGN
QIKKIEFIDFQKNEMTGTWGISTKINEQFSISFSEDRIGGKLRALGYQPNEIGFSK
20 DINSNNQ

Sequence description:

25 A] Length: 351 bp - 117 aa (Partial sequence)
B] Obvious signal peptide and Shine Dalgarno
sequence upstream of the ATG start codon.

30

ID-86

Clone ID-74

(SEQ ID NO: 21)
35 ATGTCAAATCAATATGATTATATCGTTATTGGTGGAGGTAGT
GCAGGCAGTGGTACCGCTAATAGGGCAGCCATGTATGGAGC
AAAAGTCCTGTTAATTGAAGGTGGACAAGTAGGTGGAAGTTG
TGTTAACTTAGGTTGTGTACCTAAGAAAATCATGTGGTATGG
TGCACAAGTTTCTGAGACACTCCATAAGTATAGTTCAGGTTA
40 TGGTTTTGAAGCCAATAATCTTAGTTTTGATTTTACTACTCTA
AAAGCTAATCGCGATGCTTACGTGCAGCGGTCTAGACAGTCG
TATGCCGCTAATTTTGAGCGTAATGGGGTCGAAAAGATTGAT
GGATTTGCTCGTTTTTATTGATAACCATACTATTGAAGTGAATG
GTCAGCAATATAAAGCTCCTCACATTACTATTGCAACAGGTG

GACACCCTCTTTACCCTGATATTATTGGAAGTGAACCTTGGTG
AGACTTCTGATGATTTTTTTTGGATGGGAGACCTTACCAAATTC
TATATTGATTGTTGGGGCGGGCTATATCGCGGCAGAACTTGC
TGGAGTGGTTAATGAATTAGGCGTTGAAACCCATCTTGCATT
5 TAGAAAAGACCATATTCTACGCGGATTTGATGACATGGTAAC
AAGTGAGGTTATGGCTGAAATGGAGAAATCAGGTATCTCTTT
ACATGCTAACCATGTACCTAAATCTCTTAAACGCGATGAAGG
TGGCAAGTTGATTTTTTGAAGCTGAAAATGGGAAAACGCTTGT
CGTTGATCGTGTAATATGGGCTATCGGCCGTGGACCAAATGT
10 AGACATGGGACTTGAAAATACCGATATTGTTTTAAATGATAA
AGATTATATCAAAACAGATGAATTTGAGAATACTTCTGTAGA
TGGCGTGTATGCTATTGGAGATGTTAATGGGAAAATTGCCTT
GACACCGGTAGCAATTGCAGCAGGTCGTCGCTTATCAGAAAG
ACTTTTAAATCATAAAGATAACGAAAAATTAGATTACCATAA
15 TGTACCTTCAGTTATTTTTACTCACCTGTAAATTGGGACGGTA
GGACTTTCAGAAGCAGCAGCTATCGAGCAATTTGGAAAAGAT
AATATCAAAGTCTATACATCAACTTTTACCTCTATGTATACGG
CTGTTACCAGTAATCGCCAAGCAGTTAAGATGAAGCTCATAA
CCCTAGGAAAAGAGGAAAAAAGTTATTGGGCTTCATGGTGTTG
20 GTTATGGTATTGATGAAATGATTCAAGGTTTTTTCAGTTGCTAT
CAAAATGGGGGGCTACTAAAGCAGACTTTGATGATACTGTTGC
TATTCACCCAACCTGGATCTGAGGAATTTGTTACAATGCGCTA
A
(SEQ ID NO: 22)
25 MSNQYDYIVIGGGSAGSGTANRAAMYGAKVLLIEGGQVGGTC
VNLGCVPKKIMWYGAQVSETLHKYSSGYGFEANNLSFDFTTLK
ANRDAYVQRSRQSYAANFERNGVEKIDGFARFIDNHTIEVNGQ
QYKAPHITIATGGHPLYPDIIGSELGETSDDFFGWETLPNSILIVG
AGYIAAELAGVVNELGVETHLAFRKDHILRGFDDMVTSEVMAE
30 MEKSGISLHANHVPKSLKRDEGGKLIFEANGKTLVVDRVIWAI
GRGPNVDMGLENTDIVLNDKDYIKTDEFENTSVDGVYAIGDVN
GKIALTPVAIAAGRRLSERLFNHKDNEKLDYHNVPSVIFTHPVIG
TVGLSEAAAIEQFGKDNKIKVYTSTFTSMYTAVTSNRQAVKMKLI
TLGKEEKVIGLHGVGYGIDEMIQGFSVAIKMGATKADFDDTVAI
35 HPTGSEEFVTMR*

40

Clone RS-55

(SEQ ID NO: 23)

5 ATGACAAAAAACATCTTAAAACGCTTGCCTTGGCACTTACTACAGTATCA
GTAGTGACATACAGCCAGGAGGTATATGGATTAGAAAGAGAGGAATCGGT
CAAACAAGAACAAACCCAGTCAGCTTCAGAAGATGATTGGTTCGAAGAAG
ATAATGAGAGGAAAAACAAATGTTTCTAAAGAGAATTCTACTGTTGATGAA
ACAGTTAGTGATTTATTTTCTGATGGAAATAGTAATAACTCTAGTTCTAAA
ACCGAGTCAGTGGTAAGTGACCCTAAACAAGTCCCCAAAGCAAAACCAGA
10 GGTTACACAAGAAGCAAGCAATTCTAGTAATGATGCTAGCAAAGTAGAAG
TACCAAAACAGGATACAGCTTCAAAAAAGGAAACTCTAGAAACATCAACT
TGGGAGGCAAAAGATTTCGTAAGTACTAGAGGGGATACTTTAGTAGGTTTTTCA
AAATCTGGAATTAATAAGTTATCTCAAACATCACACTTGGTTTTACCAAGT
CATGCAGCAGATGGAAGTCAATTGACACAAGTAGCTAGCTTTGCTTTTACT
15 CCAGATAAAAAGACGGCCATTGCAGAATATACAAGTAGGCTAGGAGAAA
ATGGGAAACCGAGTCGTTTAGATATTGATCAGAAGGAAATTATTGATGAG
GGAGAAATATTTAATGCTTACCAGTTGACTAAGCTTACTATTCCAAATGGT
TATAAGTCTATTGGTCAAGATGCTTTTGTGGACAATAAGAATATTGCTGAG
GTTAACCTTCCTGAGAGTCTCGAGACTATTTTCAGACTATGCTTTTGCTCACA
20 TGTCTTTAAAACAAGTAAAGTTACCAGATAACCTAAAGGTCATTGGAGAA
TTAGCTTTTTTTTGATAATCAGATTGGTGGTAAGCTTTACTTGCCACGTCCT
TGATAAAATTAGCAGAACGCGCTTTCAAATCTAATCGTATTCAAACAGTTG
AATTTTTGGGAAGTAAGCTTAAGGTTATAGGAGAAGCAAGTTTTCAAGAT
AATAATCTGAGGAATGTTATGCTTCCGGATGGACTTGAAAAAATAGAATC
25 AGAAGCTTTTACAGGAAATCCAGGAGATGAACATTACAACAATCAGGTTG
TATTGCGCACAAGGACAGGCCAAAATCCACATCAACTTGCGACTGAGAAT
ACTTACGTCAATCCGGACAAATCATTGTGGCGTGCAACACCTGATATGGAT
TATACCAAATGGTTAGAGGAAGATTTTACCTATCAAAAAAATAGTGTTACA
GGTTTTTCAAATAAAGGCTTACAAAAGGTAAGACGTAATAAAAACTTAGA
30 AATTCCAAAACAACACAATGGTATTACTATTACTGAAATTGGTGATAACGC
TTTTTCGCAATGTTGATTTTCAAAGTAAAACCTTTACGTAAATATGATTTGGA
AGAAATAAAGCTCCCCCTCAACTATTCGGAAAATAGGTGCTTTTGCTTTTCA
ATCTAATAACTTGAAATCCTTTGAAGCAAGTGAAGATTTAGAAGAGATTA
AAGAGGGAGCCTTTATGAATAATCGTATTGGAAGTCTAGACTTGAAAGAC
35 AAAGTTATCAAAATAGGTGATGCTGCTTTCCATATTAATCATATTTATGCC
ATTGTTCTTCCAGAATCTGTACAAGAAATAGGACGTTTCAGCTTTTCGACAA
AATGGTGCGCTTCACCTTATGTTTATCGGAAATAAGGTTAAAACAATTGGT
GAAATGGCTTTTTTTATCCAATAAACTGGAAAGTGTAATCTCTCTGAGCAA
AAACAATTAAAGACAATTGAGGTCCAAGCTTTTTTCGGATAATGCCCTTAGT
40 GAAGTAGTCTTACCGCCAAATTTACAGACTATTCGTGAAGAGGCTTTCAA
AGGAATCATTTGAAAGAAGTGAAGGGTTCATCTACATTATCTCAGATTACT
TTTAATGCTTTTGATCAAAATGATGGGGACAAACGCTTTGGTAAGAAAGTG
GTTGTTAGGACACATAATAATTCTCATATGTTAGCAGATGGTGAGCGTTTT
ATCATTGATCCAGATAAGCTATCTTCTACAATGGTAGACCTTGAAAAGGTT

TTAAAAATAATCGAAGGTTTAGATTACTCTACATTACGTCAGACTACTCAA
ACTCAGTTTAGAGAAATGACTACTGCAGGTAAAGCGTTGTTATCAAAATCT
AACCTCCGACAAGGAGAAAAACAAAAATTCCTTCAAGAAGCACAATTTT
CCTTGGTCGCGTTGATTTGGATAAAGCCATAGCTAAAGCTGAGAAGGCTTT
5 AGTGACCAAGAAGGCAACAAAGAATGGTCATTTGCTTGAGAGGAGTATTA
ACAAAGCGGTATTAGCTTATAATAATAGTGCTATTAAGCTAATGTTA
AGCGCTTGGAAAAAGAGTTAGACTTGCTGACAGATTTAGTCGAGGGAAAA
GGACCATTAGCGCAAGCTACAATGGTACAAGGAGTTTATTTATTAAAGAC
GCCTTTACCATTTGCCAGAATATTATATCGGATTGAACGTTTATTTTGACAA
10 GTCTGGAAAAATTGATTTATGCACTTGATATGAGTGATACTATTGGCGAGGG
ACAAAAAGATGCATATGGTAATCCTATATTAAATGTTGACGAGGATAATG
AAGGTTATCATACCTTGGCAGTTGCCACTTTAGCTGATTATGAAGGTCTTT
ATATTAAAGATATTTTAAATAGTTCCCTTGATAAGATTAAAGCAATACGCC
AGATTCCTTTGGCAAAATATCATAGATTAGGAATTTTCCAAGCTATCCGAA
15 ATGCAGCGGCAGAAGCAGACCGATTGCTTCCTAAGACACCTAAGGGGTAC
CTAAATGAAGTCCCAAATTATCGTAAAAACAAATGGAGAAAAATTTAAA
ACCAGTTGATTATAAAACGCCGATTTTTAATAAGGCTTTACCTAATGAAAA
GGTAGACGGTGATAGAGCGGCTAAAGGTCATAATATAAATGCGGAGACTA
ATAATTCTGTAGCTGTAAACACCAATAAGGTCCGAGCAGCAATTACATAAGT
20 CACAGTCTGATGTAAATTTACCTCAAACAAGTTCTAAAAATAATTTTATAT
ACGAGATTCTAGGATACGTTAGTTTATGTTTGCTTTTCCTAGTAACTGCTGG
GAAAAAAGGAAAACGAGCAAGAAAATAA

(SEQ ID NO: 24)

MTKKHLKTLALALTTVSVVTYSQEVYGLEREESVKQEQTQSASEDDWFEEDN
25 ERKTNVSKENSTVDETVDLFS DGNSNNSSSKTESVVSDPKQVPKAKPEVTQE
ASNSSNDASKVEVPKQD TASKKETLETSTWEAKDFVTRGDTLVGFSKSGINKL
SQTSHLVLP SHAADGTQLTQVASFAFTPD KKT AIAEYTSRLGENGKPSRLDIDQ
KEIIDEGEIFNAYQLTKLTIPNGYKSIGQDAFVDNKNIAEVNLPESLETISDYAF
AHMSLKQVKLPDNLKVIGELAFFDNQIGGKLYLPRHLIKLAERAFKSNRIQTV
30 EFLGSKLVIGEASFQDNNLRNVMLPDGLEKIESEAFTGNPGDEHYNNQVVLR
TRTGQNP HQ LATENTYVNPDKSLWRATPDMDYTKWLEEDFTYQKNSVTGFS
NKGLQKVRRNKNLEIPKQHNGITTEIGDNAFRNVDFQSKTLRKYDLEEIKLPS
TIRKIGAF AFQSNNLKSFEASEDLEEIKEGAFMNNRIGTLDLKD KLIKIGDAAFH
INHIYAIVLPESVQEIGRSAFRQNGALHLMFIGNKVKTIGEMAFLSNKLESVNL
35 SEQKQLKTIEVQAFSDNALSEVVLPPNLQTIREEAFKRNHLKEVKGSSTLSQITF
NAFDQNDGDKRFGKKVVVRTHNNSHMLADGERFIIDPKLSSTMVDLEKVL
KIIEGLDYSTLRQTTQTQFREMTTAGKALLSKSNLRQGEKQKFLQEAQFFLGR
VDLDKAIKAEKALVTKKATKNHLLERSINKAVLAYNNSAIKKANVKRLEK
ELDLLTDLVEGKGPLAQATMVQGVYLLKTPLPLPEYYIGLNVYFDKSGKLIYA
40 LDMSDTIGEGQKDAYGNPILNVDEDNEGYHTLA VATLADYEGLYIKDILNSSL
DKIKAIRQIPLAKYHRLGIFQAIRNAAAEADRLLPKTPKGYLNEVPNYRKKQM
EKNLKPVDYKTPIFNKALPNEKVDGDRAAKGHNINAETNNSVA VTPIRSEQQL
HKSQSDVNLPQTSSKNNFIYEILGYVSLCLLFLVTAGKKGKRARK*

Sequence description:

- 5 A] Length 3168 bp - 1056 aa (Partial sequence)
 B] Obvious signal peptide with Shine Dalgarno
 sequence upstream of the ATG start codon.

10 ID-88

Clone RS-56

15

(SEQ ID NO: 25)

GCAGGATACATCATGCACAAGCACGAGGCTATCGTGTCATGCTGGGGTCA
ACCCAGGAAGACATGTCGGCACAAGCTGAAGATTTCTTTACAGTCTGTACA
CAATAAAGAGACGGGTAAGAGCGCTTTTAATGACAAAGAACGACTAGCAA
TT

20

(SEQ ID NO: 26)

AGYIMHKHEAIVSCWGQPRKTCRHKLKISLQSVHNKETGKSAFNDKERLAI

- 25 Sequence description:

- A] Length:153 bp - 51 aa (partial sequence)
 B] No signal peptide visible, insufficient
 sequence data to determine the presence of a
30 Shine Dalgarno sequence.

35 ID-89

Clone RS-58

(SEQ ID NO: 27)

GTGTCATTTATGCAAAGAAAATCCTATTTAAAATCCATGAGTGTTCTTACT
TTAACAGCTTGTCTTATATCAGGATATGTGGTTAAAGATATTGCTATGTTA
CATGCAGTATCTGCCAGTGAGAAGAAAGCAAATAATGTCAGTCCGAGAGA
AAATCTCTACAGGGCTGTCAATGATAATTGGCTAGCCAATACAAAACCTCA
AACAAGGGGCAGACTAGTGTTAATAGTTTTTCAGAAATTGAGGATAAATTA
AAGCAACTGTTAGTGTCTGATATGGCTAAAATGGCCTCAGGAAAGATTGA

40

5 AACAACCAATGATGAACAGAAAAAATGGTTGCATACTATAAACAAGGTA
TGGACTTTAAAACAAGAGATAAAAAATGGTCTCAAACCTCTAAAACCAGTT
TTACAAAAAAGTTGAAGCAGTCTCTTCAATGAAAGACTTTCAAAGTTTGGCC
CATGATTTTGTGATGAGTGGTTTTGTTTTACCATTTGGTTTGACTGTGGAAA
CCAATGCTCGAGATAATAGCCAAAAGCAATTGGTGCTTCGTCAAGCACCC
GCATTACTTGAATCACCTGACCAATATAAGAAGGGCAATAAAGAAGGTGA
GGCTAAATTATCAGCTTACCGTACTTCAGCAATGGCTTTGCTTAAACAAGC
TGGAAAAAGTAACATTGAAGATAGAAAACTAGTTAAACAAGCTATAGCAT
TTGATAGACTCTTATCAGAAAAAACGCAAGTTGATCAAAGTAAAATCACA
10 GCTGAAAGTGAGACAGCTGCGGGGCGATATAACCCTGAAAGTATGGAAAC
GGTTCACAATTACGCCAAGGAATTTGACTTTAAAGAATTGATTGAAAAACT
AGTTGGGCCAACGAATAAGGCAGTCAATGTAGAAGATAAAACTTATTTTA
AACAGGTTAATGATGTTATAAATAGTAAACAATTAGCCAATATGAAAGCA
TGGATGATGATTTCTATGCTAGTTGATCAATCAGATTTTCTAGGAGAACAA
15 AATCGTCAAGCAGCGAGTGCTTTTAAGAATGTTGCGTCTGGTTTGACTCAG
ATTGAATCGAAAGAAAAAATGCTTACACCCAATTAG

20 (SEQ ID NO: 28)
MSFMQRKSYLKSMVLTLTACLISGYVVKDIAMLHAVSASEKKANNVSPREN
LYRAVNDNWLANTKLKQGQTSVNSFSEIEDKLKQLLVSDMAKMASGKIETTN
DEQKKMVAYYKQGMDFKTRDKNGLKPLKPVLPQKLEAVSSMKDFQSLAHDF
VMSGFVLPFGLTVETNARDNSQKQLVLRQAPALLES PDQYKKGNKEGEAKLS
AYRTSAMALLKQAGKSNIEDRKLVKQAIAFDRLLESEKTQVDQSKITAESETAA
25 GRYPESMETVHNYAKEFDFKELIEKLVGPTNKA VNVEDKTYFKQVNDVINS
KQLANMKAWMMISMLVDQSDFLGEQNRQAASAFKNVASGLTQIESKEKMLT
PN*

30 Sequence description:

35 A] Length: 1095 bp - 365 aa (full length gene)
B] an GTG (possible ATG start codon located 7 bp
further downstream) start codon with an obvious
signal peptide. Shine Dalgarno sequence present
upstream of the ORF.

40 ID-90

Clone RS-59

(SEQ ID NO: 29)

ATGGAAATGCCTAAAAGAAATGAATTACTCAATAAAGAAATTAAAATGAG
TATTGATAAACTTAGATATAAAGAACCAGAGAGTGAACATGACAAGCGAC
CTACTTTTTATTTGGTAGTACTTATACTTGTTACTGTAGCAGTTATATTGTC
GTTATTTAAATATTTTTTATAG

5

(SEQ ID NO: 30)

MEMPKRNELLNKEIKMSIDKLRYKEPESEHDKRPTFYLVVLILVTVAVILSLFK
YFL*

10

Sequence description:

A] Length: 174 bp - 58 aa(full length gene)

B] No obvious signal peptide, but Shine

Dalgarno sequence is present upstream of ATG
start codon.

15

20

ID-91

Clone RS-62 (partial sequence)

25

(SEQ ID NO: 31)

ATGCAGGTATTTTTAAATATTGTCAATAAATTCTTTGATCCAGTTATTCATA
TGGGTTTCGGGAGTTGTGATGCTAATTGTCATGACAGGTTTAGCCATGATAT
TTGGAGTGAAGTTTTCTAAAGCACTTGAAGGTGGTAT

30

(SEQ ID NO: 32)

MQVFLNIVNKFDPVIHMGSGVVMLIVMTGLAMIFGVKFSKALEGG

Sequence description:

35

A] Length: 141 bp - 41 aa (partial sequence

B] Shine Dalgarno sequence present upstream of
ATG start codon with a possible signal peptide
present

40

ID-92

Clone RS-69 (partial sequence)

(SEQ ID NO: 33)

5 ATGAAAAAGAAAACATTCAGTGCTTATAACTTTTTAACGGCTCTTATCCTT
TGTCTTTTGACAGTGCTTTTTATCTTTCCATTTTATTGGATTATGACAGGAG
CTTTTAA

(SEQ ID NO: 34)

MKKKTFSAYNFLTALILCLLTVLFIFPFYWIMTGAF

10

Sequence description:

A] Length: 110 bp -36 aa (Partial sequence)

15

B] Possible signal peptide with Shine Dalgarno
sequence directly upstream of the ATG start
codon.

20

ID-93

Clone RS-70

25

(SEQ ID NO: 35)

30 ATGACTGAGAACTGGTTACATACTAAAGATGGTTCAGATATTTATTATCGT
GTCGTTGGTCAAGGTCAACCGATTGTTTTTTTACATGGCAATAGCTTAAGT
AGTCGCTATTTTGATAAGCAAATAGCATATTTTCTAAGTATTACCAAGTT
ATTGTTATGGATAGTAGAGGGCATGGCAAAAGTCATGCAAAGCTAAATAC
CATTAGTTTCAGGCAAATAGCAGTTGACTTAAAGGATATCTTAGTTTCATTT
AGAGATTGATAAAGTTATATTGGTAGGCCATAGCGATGGTGCTAATTTAGC
TTTAGTTTTTCAAACGATGTTTCCAGATATGGTTAGAGGGGCTTTTGCTTAAT
35 TCAGGGAACCTGACTATTCATGGTCAGCGATGGTGGGATATTCTTTTAGTA
AGGATTGCCTATAAATTCCTTCACTATTTAGGGGAAACTCTTCCGTATATG
AGGCAAAAAGCTCAAGTTATTTTCGCTTATGTTGGAGGATTTGAAGATTAGT
CCAGCTGATTTACAGCATGTGTCAACTCCTGTAATGGTTTTGGTTGGAAAT
AAGGACATAATTAAGTTAAATCATTCTAAGAAACTTGCTTCTTATTTTCCA
AGGGGGGAGTTTTATTCTTTAGTTGGCTTTGGGCATCACATTATTAAGCAA
40 GATTCCCATGTTTTTAATATTATTGCAAAAAAGTTTATCAACGATACGTTG
AAAGGAGAAATTGTTGAAAAAGCTAATTGA

(SEQ ID NO: 36)

MTENWLHTKDGSDIYYRVVGQGQPIVFLHGNSLSSRYFDKQIAYFSKYYYQVIV
MDSRGHGKSHAKLNTISFRQIAVDLKDILVHLEIDKVILVGHSDGANLALVFQ

TMFPDMVRGLLLNSGNLTIHGQRWWDILLVRIA YKFLHYLGKLFYMRQKA
QVISLMLEDLKISPADLQHVSTPVMVLVGNKDIIKLNHKKLASYPFRGEFYSL
VGFGHHIIKQDSHVFNIIAKKFINDTLKGEIVEKAN*

5

Sequence description:

- 10 A] Length: 744 bp - 248 aa (full length gene)
 B] No obvious signal peptide, but Shine
 Dalgarno sequence upstream of the ATG start
 codon.

15

ID-94

20 Clone RS-71

 (SEQ ID NO: 37)
ATGGTAGCAAAAGAGTTAGGTAAAAATAGCTTTACTATCCCAACTATTTGT
TCTAATTGCTCCGCAGGTACTGCCATTGCAGTTGTATATAATGATGACCAT
TCTTTCTTAAGATACGGCTATCCCGAGTCTCCACTTCATATTTTTATCAATA
25 CACGGATCATTGCACAGGCACCAAGCAAATATTTTTGGGCTGGTATTGGGG
 ACGGTATTTCAAAAGCCCCTGAAGTAGAACGTGCTACCTTAGAGGCTAAG
 ACCAATAAACTACCACATACTGCAGTGTTAGGACAAGCAGTCGCTCTGTCT
 TCAAAGGAAGCTTTTTATCAATTTGGTGAACAAGGTCTAAAAGACGTTGAA
 GCTAATTTAGCTTCGCGTGCAGTTGAAGAAATTGCGCTTGATATCTTA

30

 (SEQ ID NO: 38)
MVAKELGKNSFTIPTICSNCSAGTAIAVVYND DHSFLRYGY P ESPLHIFINTRIILA
QAPSKYFWAGIGDGISKAPEVERATLEAKTNKLPHTAVLGQAVALS SKEAFY
QFGEQGLKDVEANLASRAVEEIALDIL

35

Sequence description:

- 40 A] Length: 405 bp - 135 aa (Partial sequence)
 B] No obvious Shine Dalgarno sequence upstream
 of the ATG start codon, probable signal
 peptide present at the N-terminus.

ID-95

Clone RS-73

5 (SEQ ID NO: 39)
TTGAGGGAACTTACTGGAAAATTTCAAGCGATTGCGATAAAATAAATCTT
GCAGAGTTTTCTAGAGAAAGGAGGTCAGATTTATTGGAGTGGCAAGATCT
AGCGCAGTTACCTGTATCTATTTTTTAAAGACTATGTTACAGATGCTCAAGA
CGCGGAAAAACCTTTTATATGGACAGAAGTATTTTTTAAGGGAGATTAATCG
CTCAAATCAAGAAATTATTTTGCATATTTGGCCGATGACTAAGACAGTCAT
10 TCTGGGGATGTTAGATCGAGAATTACCACATTTAGAATTAGCTAAAAAAG
AAATCATCAGTCGTGGTTATGAACCAGTTGTTTCGGAATTTTGGAGGTCTCG
CAGTTGTAGCTGATGAAGGAATTTTAAATTTTTCATTGGTTATTCCAGATGT
TTTTGAGAGAAAATTGTCTATCTCAGATGGGTATCTTATAATGGTCGATTTT
ATTAGAAGTATATTTTCGGATTTTTATCAACCTATTGAGCACTTTGAAGTA
15 GAGACCTCCTATTGTCCTGGTAAGTTTGATCTTAGTATAAATGGCAAAAAA
TTTGCTGGCTTGGCTCAGCGCCGTATAAAGAATGGTATTGCGGTATCAATT
TACCTTAGCGTTTGTGGCGATCAAAAAGGGCGGAGTCAAATGATTTCAGAT
TTTTATAAGATTGGTCTAGGTGATACGGGTAGTCCAATTGCTTATCCAAAT
GTAGATCCTGAAATTATGGCTAATCTATCTGATCTATTAGATTGTCCTATG
20 ACAGTAGAAGATGTTATTGATCGTATGTTGATTAGCCTTAAACAAGTAGGT
TTTAATGATCGTTTACTGATGATTAGACCCGATTTAGTTGCAGAGTTTGAT
AGATTTTCAGGCTAAGTCTATGGCTAATAAGGGGGATGGTGAGCAGAGATGA
ATAA
25 (SEQ ID NO: 40)
MRETYWKISSDCDKINLAEFSRERRSDLLEWQDLAQLPVSIFKDYVTDAQDAE
KPFIWTEVFLREINRSNQEILHIWPMTKTVILGMLDRELPHLELAKKEIISRGYE
PVVRNFGGLAVVADEGILNFSLVIPDVFERKLSISDGYLIMVDFIRSIFSDFYQPI
EHFEVETSYCPGKFDLSINGKKFAGLAQRRIKNGIAVSIYLSVCGDQKGRSQMI
30 SDFYKIGLGD TGSP IAYPNVDPEIMANLSDLLDCPMTVEDVIDRMLISLKQVGF
NDRLLMIRPDLVAEFDRFQAKSMANKGMVSRDE*

Sequence description:

35 A] Length: 921 bp -307 aa (Full-length gene sequence)
B] No obvious Shine Dalgarno sequence upstream
of the TTG start codon or signal peptide
visible: Actual start point may be a further
40 85 bp downstream (TTG). This start point is
preceded by a typical Shine-Dalgarno sequence.

ID-96

Clone RS-74

5

(SEQ ID NO: 41)

TTGGAAGGTTTACTTATTGCATTGATTCCCATGTTTGCGTGGGAAAGTATT
GGATTTGTTAGTAATAAAAATTGGAGGGCGTCCAAATCAACAAACATTTGG
AATGACTTTAGGAGCATTGCTATTTGCGATTATCGTATGGTTATTTAAACA
10 GCCAGAGATGACTGCCTCATTGTGGATTTTTTGGTATCTTAGGTGGTATCCT
ATGGTCAGTCGGCCAAAATGGTCAATTTCAAGCAATGAAATATATGGGAG
TCTCTGTTGCTAATCCACTGTCAAGTGGTGCACAATTAGTAGGTGGAAGCC
TAGTTGGTGCTTTAGTCTTTCATGAATGGACTAAGCCAATCCAATTTATTTT
AGGATTGACAGCGTTGACATTATTAGTTATCGGCTTCTATTTCTCAAGTAA
15 ACGTGATGTTTCAGAACAAGCTTTGGCAACACATCAAGAGTTTTCAAAG
GATTTGCTACAATTGCTTATTCAACTGTAGGTTACATCTCGTACGCAGTTTT
ATTTAACAACATTATGAAGTTCGACGCTATGGCCGTCATTTTACCCATGGC
TGTTGGAATGTGTCTAGGTGCAATTTGTTTCATGAAGTTTCGTGTAACTTT
GAGGCTGTTGTTGTTAAAAATATGATTACAGGTCTCATGTGGGGCGTTGGT
20 AATGTCTTCATGTTATTGGCAGCAGCTAAAGCAGGGGCTAGCAATTGCTTTT
AGTTTTTCTCAACTTGGAGTAATTATCTCTATTATTGGTGGTATTTTATTTT
AGGTGAGACAAAAACGAAGAAAGAGCAGAAATGGGTTGTCATGGGTATC
CTTTGTTTTGTTATGGGTGCTATATTACTTGGTATTGTTAAATCTTATTAA

(SEQ ID NO: 42)

25 MEGLLIAIPMFAWESIGFVSNKIGGRPNQQTFGMTLGALLFAIIVWLFKQPEM
TASLWIFGILGGILWSVGQNGQFQAMKYMGSVSVANPLSSGAQLVGGSLVGAL
VFHEWTKPIQFILGLTALTLLVIGFYFSSKRDVSEQALATHQEFSKGFATIAYST
VGYSISYAVLFNNIMKFDAMAVILPMAVGMCLGAICFMKFRVNFEAVVVKNMI
TGLMWGVGNVFMLLAAAKAGLAIAFSFSQLGVIISIIGGILFLGETKTKKEQK
30 WVVMGILCFVMGAILLGIVKSY*

Sequence description:

35

A] Length: 867 bp - 289 aa (full-length gene)
B] Possible Shine Dalgarno sequence upstream of
GTG start codon, no obvious signal peptide
present.

40

ID-97

Clone RS-75

(SEQ ID NO: 43)
5 ATGACAACCTTACTACGAAGCTATAAACTGGAACGAAATTGAAGATGTTAT
TGATAAATCAACTTGGGAAAACTAACCGAACAATTTTGGCTCGATACAC
GTATCCCTTTATCAAATGACTTAGACGATTGGCGCAAACCTTCCGCTCAAG
AAAAAGATCTTGTTGGCAAGGTTTTTGGAGGCTTAACCCTACTTGATACCA
TGCAATCAGAACTGGTGTGGAAGCTATTCGTGCCGATGTTTCGCACGCCTC
ACGAAGAAGCTGTCTTAAACAATATTCAATTCATGGAATCTGTTACGCTA
10 AATCTTATTCTTCAATTTTCTCAACTTTAAATACTAAATCAGAAATTGAAG
AAATTTTCGAGTGGACTAATAATAATGAGTTCCTTCAAGAAAAAGCACGT
ATTATCAATGACATTTATGCTAATGGAAATGCCCTTCAAAAAAAGGTGGCT
TCCACCTACCTCGAACTTTCCTTTTTTATTCTGGCTTTTTCACACCTCTTA
CTATTTGGGAAATAATAAGTTAGCAAATGTTGCTGAAATCATTAAATTAAT
15 TATTCGTGATGAATCTGTACATGGTACTTATATCGGTTACAAATTCCAGCTT
GGTTTTAACGAATTACCAGAAGATGAGCAAGAGAATTTTCGTGATTGGAT
GTATGACCTCCTTTATCAGCTGTATGAAAACGAAGAAAAATACACCAAGA
CACTTTATGATGGCGTAGGATGGACTGAAGAAGTTATGACCTTTTTTACGCT
ACAATGCTAATAAAGCTCTTATGAATTTAGGACAAGATCCTTTATTCCCAG
20 ATACAGCAAATGATGTCAACCCAATTGTTATGAATGGTATTTCAACAGGAA
CATCAAACCATGACTTCTTCTCTCAAGTAGGTAATGGTTACCTACTTGGTA
GCGTTGAAGCTATGCATGATGACTATAACTATGGATTATAA
(SEQ ID NO: 44)
25 MTTYEAINWNEIEDVIDKSTWEKLTEQFWLDTRIPLSNDLDDWRKLSAQEK
DLVGKVFGGLTLLDTMQSETGVEAIRADV RTPHEEAVLNNIQFMESVHAKSY
SSIFSTLNTKSEIEEIFEWTNNNEFLQEKARIINDIYANGNALQKKVASTYLETF
LFYSGFFTPLYLGNKLANVAEIKLIIRDES VHGTYIGYKFQLGFNELPEDEQ
ENFRDWMYDLLYQLYENEEKYTKTLYDGVGWTEEVMTFLRYNANKALMNL
30 GQDPLFPDTANDVNPIVMNGISTGTSNHDFFSQVGNGYLLGSVEAMHDDYN
YGL*

Sequence description:

35 A] Length: 960 bp - 320 aa (full length gene)
B] Shine Dalgarno sequence present upstream of
ATG start codon, but no signal peptide
present.

40

ID-98

Clone RS-77 (partial sequence)

(SEQ ID NO: 45)
5 ATGAATTGGTCACGTATCTGGGAACTCGTAAAAATTAATATCCTTTATTCA
AACCCCTCAGACTCTATCGGCACTAAGAAAAAAGCAAGAAAAGCATCCTAA
AAAAGAATTTTCAGCTTATAAATCCATGTTTAGAAATCAGTTATTTTCAGAT
TTTGCTCTTTTCAATAATTTATGTATTTCTCTTTGTATCACTTGATTTTAAAG
AATATCCGGGCTATTTACGTTCTACATTGGTATCTTTACACTAGTATCCAT
10 TATCTACTCTTTTATTGCGATGTACAGTGTTTTCTATGAGAGTGACGATGTT
AA

(SEQ ID NO: 46)
MNWSRIWELVKINILYSNPQTLSALRKKQEKHPKKEFSAYKSMFRNQLFQILL
FSIIYVFLFVSLDFKEYPGYFTFYIGIFTLVSIISFIAMYSVFYESDDV

15
Sequence description:

20 A] Length: 311 bp - 103 aa (Partial sequence)
B] Shine Dalgarno sequence present upstream of
ATG start codon, no obvious signal peptide at
N-terminus.

ID-99

25
Clone RS-78 (partial sequence)

(SEQ ID NO: 47)
30 TAATCTTTTAGTCAACGGAGCAACAGGAAAATTGCAGGCTATGCGACAGA
TATTCCACCACATAATTTAGCAGAAGTCATTGATGCTGTCGTGTACATGAT
TGATCACCCTAAAGCTAAATTAGATAAATTAATGGAATTTCTACCTGGTCC
AGATTTTCCAACCTGGCGCTATCATTCAAGGAAAAGATGAAATTCGTAAGG
CATATGAGACTGGTAAGGGGAGAGTAGCGGTTTCGCTCGCGAACTGCTATT
GAAACCTTAAAAGGTGGTAAGAAACAAATTATTGTTACTGAAATTCCTTAT
35 GAAGTTAAT

(SEQ ID NO: 48)
SFSQRSNRKIAGYATDIPPHNLAEVIDAVVYMIDHPKAKLDKLMEFLPGPDFPT
GAIQGKDEIRKAYETGKGRVAVRSRTAETLKGGKKQIIVTEIPYEVN

40
Sequence description:

A] Length: 312 bp - 104 aa (Partial sequence)
B] No obvious Shine Dalgarno sequence or a

signal peptide. Both N- and C- termini of ORF yet to be elucidated.

5 ID-100

Clone RS-79

10 (SEQ ID NO: 49)
ATGGGACGTAAGTGGGGCCAATATTGTTGCCAAAAAGACTGCTAAAGATGG
TGCTAACTCAAAAGTATACGCTAAATTCGGTGTTGAAATATATGTTGCTGC
AAAGCAAGGTGAACCAGACCCCGAGTCAAACCTCAGCTCTAAAATTCGTTT
TGGACCGTGCTAAGCAAGCACAAAGTTCCAAAGCATGTTATTGATAAAGCG
15 ATTGATAAAGCCAAAGGAAACACAGATGAAACTTTCGTAGAGGGGACGCTA
TGAAGGTTTTTGGTCCAAATGGTTCAATGATTATTGTGGATACTTTGACATC
AAATGTTAACCGTACGGCAGCAAATGTACGTACTGCTTACGGTAAGAACG
GTGGCAATATGGGAGCTTCAGGATCGGTATCCTACTTATTTGATAAAAAAG
GTGTCATCGTTTTTTGCTGGTGATGATGCTGACACTGTCTTCGAACAATTACT
20 TGAAGCGGATGTAGACGTAGATGATGTTGAAGCAGAAGAGGGAACAATA
ACAGTTTATACCGCCCCAACAGATCTTCATAAAGGTATCCAAGCACTTCGC
GATAATGGTGTAGAAGAATTCCAAGTTACTGAACTTGAAATGATTCCTCAA
TCAGAAGTAGTATTGGAAGGTGATGACCTTGAAACTTTTGAAAAGCTT
25 (SEQ ID NO: 50)
MGRKWANIVAKKTAKDGANSKVYAKFGVEIYVAAKQGEPPDPESNSALKFVL
DRAKQAQVPKHVIDKAIDKAKGNTDETFVEGRYEGFGPNGSMIIVDTLTSNV
NRTAANVRTAYGKNNGGNMGASGSVSYLFDKKGVIVFAGDDADTVFEQLLEA
DVDVDDVEAEEGTITVYTAPTDLHKGIQALRDNGVEEFQVTELEMIPQSEVVL
EGDDLETFEKL

30

Sequence description:

35 A] Length: 654 bp - 218 aa (Partial sequence)
B] Possible Shine Dalgarno sequence upstream
of ATG start, no obvious signal peptide

40 ID-101

Clone RS-80

(SEQ ID NO: 51)

TTGGAGAAATATTTGAAGAACCCGATTACATGGATTGGATTAGTTCTTGTG
GTTACGTGGTTTTTAACTAAAAGTAGTGAATTTTGGATTGTTGGTGTGTGTG
TCTTGTTGTTAGTATTTGCTAGTCAAAGTGAT

5

(SEQ ID NO: 52)

MEKYLKNPITWIGLVLVVTWFLTKSSEFLIFGVCVLLLVFASQSD

Sequence description:

10

A] Length: 135 bp - 45 aa (partial sequence)
B] Shine Dalgarno sequence upstream of TTG
start codon with possible signal peptide
evident at N-terminus.

15

ID-102

20

Clone RS-81

(SEQ ID NO: 53)

ATGACACAATCAGATGCATATCTCTCGTTGAACGCGAAGACACGCTTTAGA
GATCGCACAGGTAATTATCATTTTACTTCGGATAAAGAGGCTGTTGAACAA
TATATGATAGAACATGTTGAACCTAATACGATGGTGTTTACATCACTAATT
GAAAAGCTAGATTATTTGGTTTCTAATAACTACTATGAATCGGACCTTCTA
AAACAATATAACCTTGAGTTTATTTGCCAAATTTTGGAGCATGCATACGCT
AAGAAATTTGCTTTTCTAAATTTTATGGGGGCTTTAAAATTTTATAATGCTT
ATGCTCTTAAT

30

(SEQ ID NO: 54)

MTQSDAYLSLNAKTRFRDRTGNYHFTSDKEAVEQYMIEHVEPNTMVFTSLIE
KLDYLVSNYYYESDLLKQYNLEFICQIFEHAYAKKFAFLNFMGALKFYNAYA
LN

35

Sequence description:

40

A] Length: 318 bp - 106 aa (Partial sequence)
B] Shine Dalgarno sequence present upstream of
ATG start codon, no obvious signal peptide

ID-103

Clone 2-11A

5

(SEQ ID NO: 55)

ATGGTATTTATGGCAAATAAGAAAAAAACAAAAGGAAAGAAAACCAGAA
GACCTACTAAGGCAGAAATAGAGCGTCAAAGAGCTATTCAAAGGATGATT
ACTGCTCTTGTTTTAACAATTATTCTCTTCTTTGGTATTATCAGATTAGGTA
10 TTTTGGTATTACAGTCTATAACGTCATCCGTTTTATGGTAGGTAGCTTGGC
TTACTTATTTATTGCGGCAACTTTAATCTACCTTTATTTCTTTAAATGGTTG
CGAAAGAAAGATAGCTTAGTAGCAGGTTTTTTGATAGCTTCTTTAGGATTA
TTGATTGAGTGGCATGCTTACCTTTTCTCAATGCCTATTTTGAAAGATAAA
GAAATTTTGCGTTCAACTGCTCGATTAATTGTGTCTGATTTAATGCAATTTA
15 AAATCACTGTTTTTGCCGGTGGAGGTATGTTGGGTGCTTTGATTTACAAGC
CAATTGCTTTTCTCTTTTCTAATATTGGTGCCTATATGATTGGTGTTCTCTTC
ATCATTTTGGGTCTCTTTTAAATGAGTTCTCTGGAAGTTTATGACATCGTCG
AATTTATTAGAGCTTTTAAAAATAAAGTGGCAGAGAAGCACGAGCAAAAT
AAAAAGGAGCGTTTTTGCTAAGCGAGAGATGAAAAAAGCAATCGCTGAACA
20 AGAGCGCATAGAGCGTCAAAAAGCTGAAGAAGAAGCTTATTTAGCTTCGG
TTAATGTAGACCCTGAAACGGGTGAGATTCTAGAGGATCAAGCTGAGGAC
AATTTGGATGATGCGCTACCACCTGAGGTAAGTGAAACATCAACTCCGGT
ATTTGAGCCAGAGATCCTTGCTTATGAGACATCGCCTCAAAATGATCCTTT
ACCAGTAGAGCCGACAATTTATTTAGAAGACTATGATTTCGCCGATTCTCTAA
25 TATGAGAGAAAATGATGAGGAAATGGTTTATGATTTAGATGATGATGTAG
ATGATAGTGATATAGAAAATGTCGACTTTACACCTAAAACGACACTGGTTT
ATAAATTACCAACGATAGATTTATTTGCACCAGATAAGCCTAAAAATCAAT
CCAAAGAAAAGGATTTAGTCCGAAAGAATATCAGAGTTTTAGAAGAAACA
TTTAGAAGTTTTTGGTATCGATGTAAAAGTAGAACGTGCTGAAATTGGACCA
30 TCAGTTACTAAATATGAAATTAAACCAGCAGTTGGAGTTCGTGTGAATCGT
ATTTCAAATCTATCTGACGACCTAGCTCTTGCTCTTGACAGCAAAAGATGTG
CGTATAGAAGCACCAATTCCTGGAAAATCATTAAATAGGTATTGAAGTTCCT
AACTCAGAAATTGCAACGGTTTTCTTTCCGCGAACTTTGGGAACAATCTGAT
GCCAATCCTGAAAACCTTTTAGAAGTACCACTAGGAAAAGCTGTTAACGG
35 CAATGCTCGCAGTTTTTAACTTAGCTAGAATGCCGCATCTTTTGGTAGCTGG
TTCAACTGGTTCAGGTAAATCTGTGGCAGTTAATGGAATTATTTCAAGTAT
TTTGATGAAGGCACGTCCAGATCAAGTTAAGTTTATGATGATTGATCCCAA
AATGGTTGAATTATCTGTTTATAATGATATTCCACATTTATTAATCCCTGTT
GTAACCAATCCGCGTAAAGCAAGTAAGGCACTCCAAAAAGTTGTTGATGA
40 AATGGAAAATCGATACGAGTTATTTAGCAAAATTGGTGTGCGTAATATAG
CAGGTTATAATACAAAGGTTGAAGAGTTTAATGCTTCCTCTGAGCAAAAAC
AAATGCCTTTGCCTTTAATCGTTGTCATTGTAGATGAATTGGCTGACTTGAT
GATGGTTGCTAGTAAAGAAGTTGAAGATGCTATTATTCGTTTGGGGGCAAAA
AGCACGTGCTGCAGGTATCCATATGATTCTTGCAACTCAACGTCCATCCGT

AGATGTTATTTCTGGTTTGATTAAAGCAAATGTTCCGTCGCGTATTGCATTT
GCTGTTTCAAGTGGTACTGATAGCCGTACGATCCTTGATGAAAATGGTGCT
GAAAAGCTCTTGGGACGGGGTGACATGCTCTTTAAGCCTATTGATGAGAAT
CATCCAGTACGACTACAAGGTTTCCTTTATTTTCAGATGATGATGTTGAAAGG
5 ATCGTTGGTTTTATCAAAGACCAAGCCGAGGCTGACTATGATGATGCCTTT
GATCCTGGAGAAGTATCTGAAACAGATAACGGCTCTGGTGGTGGCGGCGG
AGTACCTGAAAGTGATCCTCTTTTTTGAAGAAGCCAAGGGACTCGTTTTAGA
GACGCAAAAAGCAAGTGCCTCAATGATTCAACGCCGATTGTCTGTTGGTTT
CAATAGAGCAACAAGACTAATGGAAGAATTAGAAGCAGCGGGGGGTTATTG
10 GTCCAGCAGAAGGAACCAAGCCACGAAAAGTTTTAATGACTCCAACCTCCG
AGTGAATAA

(SEQ ID NO: 56)
MVFMANKKKTKGKKTRRPTKAEIERQRAIQRMITALVLTILFFGIIRLGIFGIT
VYNVIRFMVGSLAYLFIAATLIYLYFFKWLRKKDSL VAGFLIASLGLLIEWHA
15 YLFSMPILKDKEILRSTARLIVSDLMQFKITVFAGGGMLGALIYKPIAFLFSNIG
AYMIGVLFILGLFLMSSLEVYDIVEFIRAFKNKVAEKHEQNKKERFAKREMK
KAIAEQERIERQKAE EEA YLASVNVD PETGEILEDQAEDNLDDALPPEVSETST
PVFEPEILAYETSPQNDPLPVEPTIYLEDYDSPIPNMRENDEEMVYDLDDD VDD
SDIENVDFTPKTTLVYKLPTIDLFAPDKPKNQSKEKDLVRKNIRVLEETFRSFGI
20 DVKVERAEIGPSVTKEYEIKPAVGVRVNRISNLSDDLALALAAKDVRIEAPIPGK
SLIGIEVPNSEIATVSFRELWEQSDANPENLLEVPLGKAVNGNARSFNLARMPH
LLVAGSTGSGKSVAVNGISSLMKARPDQVKFMMIDPKMVELSVYNDIPHLLI
PVVTNPRKASKALQKVVD MENRYELFSKIGVRNIAGYNTKVEEFNASSEQK
QMPLPLIVVIVDELADLMMVASKEVEDAIIRLGQKARAAGIHMILATQRPSVD
25 VISGLIKANVPSRIAFAVSSGTD SRTILDENGA EKLLGRGDMLFKPIDENHPVRL
QGSFISDDDVERIVGFIKDQAEADYDDAFDPGEVSETDNGSGGGGGVPESDPL
FEEAKGLVLETQKASASMIQRRLSVGFNRATRLMEELEAAGVIGPAEGTKPRK
VLMTPTPSE*

30

Sequence description:

35 A] Length: 2451 bp - 817 aa (Full-length gene)
B] Shine Dalgarno sequence present upstream of
ATG start codon, possesses a potential signal
peptide

40 ID-104

Clone 2-18/22b

(SEQ ID NO: 57)

ATGTCACAAGAGCAAGGAAAAATTTATATTGTAGAAGATGATATGACGAT
TGTGTCACTTTTAAAAGATCATTTATCAGCTAGCTATCATGTCTCTAGTGTC
AGCAATTTTCGTGATGTGAAACAAGAAATTATCGCATTTC AACCCGATTTG
ATACTAATGGATATTACGTTACCCTATTTTAATGGTTTTTACTGGACTGCAG
5 AATTGCGTAAGTTTTTAACAATTCCTATTATTTTCATTTCATCTAGTAATGA
TGAAATGGATATGGTTATGGCATTAAATATGGGGGGGTGATGACTTTATTTT
AAAACCATTCTCTCTAGCTGTATTAGATGCTAAGCTAACTGCTATTTTAAG
GAGAAGTCAACAATTTATCCAACAGGAATTA ACTTTTGGGGGGATTTACGTT
GACAAGAGAAGGGTTATTGTCTAGCCAAGATAAAGAGGTTATTTTATCGC
10 CAACAGAAAATAAAAATCCTATCTATCTTGCTCATGCATCCTAAACAAGTAG
TCTCAAAAGAGTCTCTATTAGAGAACTTTGGGAAAATGATAGTTTTTATTG
ATCAAAATACACTTAATGTTAATATGACACGCTTACGTAAAAAAATTGTCC
CAATAGGTTTTGATTACATTCATACAGTGAGAGGAGTTGGGTATTTACTAC
AATGA

(SEQ ID NO: 58)

MSQEQGKIYIVEDDMTIVSLLKDHLASASYHVSSVSNFRDVKQEIIAFQPDILM
DITLPYFNGFYWTAELRKFLTPIIFISSNDEMMDMVMALNMGGDDFISKPFSLA
VLD AKLTAILRRSQQFIQQELTFGGFTLTREGLLSSQDKEVILSPTENKILSILLM
20 HPKQVVSKESSLLEKLWENDSFIDQNTLNVNMTRLRKKIVPIGFDYIHTVRGVG
YLLQ*

Sequence description:

A] Length: 669 bp - 223 aa (full-length gene
sequence)

B] Shine Dalgarno sequence present upstream of a GTG start codon.
Was not identified directly by LEEP. This gene was found upstream of
30 gene ID-10 described in WO 00/06736.

ID-105

Clone 2-20

(SEQ ID NO: 59)

ATGTATCAAACCTCAGACAAATAAGGAAAAATTTGTTTTATTTTTGAAATTA
TTTATCCCAGTATTGATTTATCAATTTGCTAATTTTTCAGCTACTTTTATTGA
40 TTCGGTTATGACTGGACAGTATAGTCAGCTACATTTGGCAGGTGTGTCAAC
TGCTAGTAATTTATGGACTCCGTTTTTCGCTTTATTAGTAGGTATGATTTCA
GCATTAGTACCAGTAGTTGGTCAACATTTGGGTAGAGGAAATAAAGAACA
AATTCGCACAGAATTTTCATCAATTTCTATATTTAGGTTTGATACTGTCCTTA
ATATTATTTTAAATCATGCAATTTATTGCTCAACCTGTCTTGGGGAGTTTGG

5 GTTTAGAAGATGAAGTTCTAGCAGTTGGTCGTGGTTATTTAAATTATATGT
TGATTGGAATCATGCCGCTGGTGTTGTTTAGCATTTGCCGTTTCATTCTTTGA
TGCATTGGGGTTAACAAGGTTATCTATGTATCTGATGCTTTTAATTCTACCC
TTTAATTCATTTTTTTAATTATATGCTTATCTACGGTAAATTTGGTATGCCTA
GACTAGGAGGTGCGGGGGCAGGTCTTGGAACCTCTTTAACTTATTGGGCTA
TTTTTATTGGTATTATTATTGTGATGTCACCTCATCCTCAAATTAACACATA
TCATATATGGACTCTGGAAAGAATAAAAGCTCCTTTGATTATTGAAGATAT
TCGATTGGGATTACCGATTGGTTTACAAATTTTTGCAGAAGTTGCAATTTTT
GCAGTAGTAGGCTTATTCATGGCAAATTTTCTTCAATCATTATTGCAGCA
10 CATCAGGCTGCTATGAATTTTTTCATCATTAAATGTATGCATTTCTTTAAGTA
TTTCCACTGCTCTAGCTATTACAATATCGTTTGAAGTAGGGGCAGAGCGCT
TTCAGGACGCAACCACTTATAGTAGGATAGGACGCTTAACAGCGGTAGGG
ATTACATCAGGAACCTTACTATTTTTATTTCTATTTTCGTGAGAATGTAGCAG
CAATGTATAATAGTGCCCCTCACTTTGTCGCTATTACAGCTCAATTCCTAAC
15 TTATAGTCTCTTTTTCCAGTTTGCAGATGCTTATGCAGCTCCTGTACAGGGG
ATTTTACGAGGCTATAAGGATACAACAAAACCATTTATGATCGGTGCGGG
CTCTTATTGGTTATGTGCTTTGCCATTAGCGGTTATCTTAGAAAAAATAG
CCAGTTAGGTCCGTTTGCCTATTGGATTGGTTTAATCACAGGTATTTTTTGT
TGTGGTCTATTTCTAAACCAACGTCTGCAAAAGATTAAGAAGTTGTATTAT
20 TAA

(Seq ID NO: 60)
MYQTQTNKEKFVLFLKLFIPVLIYQFANFSATFIDSVMTGQYSQLHLAGVSTAS
NLWTPFFALLVGMISALVPVVGQHLGRGNKEQIRTEFHQFLYLGLILSLILFLI
MQFIAQPVLGSLGLEDEVLA VGRGYLNYMLIGIMPLVLFSICRSFFDALGLTRL
25 SMYLMLLILPFNSFFNYMLIYGKFGMPRLGGAGAGLGTSLTYWAIFIGIIIVMS
LHPQIKTYHIWTLERIKAPLIIEDIRLGLPIGLQIFAEVAIFAVVGLFMAKFSSIIA
AHQAAMNFSRLMYAFPLSISTALAITISFEVGAERFQDATTYSRIGRLTAVGITS
GTLFLFLFRENVAAAMYNAPHFVAITAQFLTYSLFFQFADAYAAPVQGILRG
YKDTTKPFMIGAGSYWLCALPLAVILEKNSQLGPFA YWIGLITGIFVCGLFLNQ
30 RLQKIKKLYY*

35 Sequence description:

- A] Length: 1341 bp - 447 aa (full length gene)
B] Shine-Dalgarno sequence present upstream of
ATG start codon, There is a potential signal
peptide sequence

40

ID-106

Clone 2-4A

(SEQ ID NO: 61)

5 TTGCTAGTTTCTTCTCTAGTTTCTTGTTCAATTTTTCTTGTCATTTTCGTCGTT
GTCTTCATCAACACGAAATAAGTCTATAAACTTATCAAATAATTTCATAGA
CTTATTATATCAATTTTCAATAAAATGCTATAATAAAACCATGTCATTTTCA
TTAAAAATTAGAAATCCATACGGTGAACATAACCGTTAAAGAACTCCTTGA
AGATTATTTTTTGATTCCACGTAAGATTAGACATTTTTTGC GTGTTAAAAAA
10 CATGTACTTATAAACAATGAATTCATTAATTGGCAAACCTGTCGTCCAAGAA
AACGATACTATTACCTTAATCTTTGATGATGAGGATTACCCTACTAAAAAA
ATTCCTCTGGGCAGAGCAGAGCTTATTGATTGTCTTTATGAGGATGAACAT
CTTATTATCGTTAATAAACCTGAAGGTATGAAAACCTCACGGTAACCAACCA
AATGAAATAGCACTGTAAATCATGTATCTGCCTATTCTGGACAAACATGC
15 TATGTTGTTTCATCGCCTAGATATGGAGACCAGTGGAGCTGTTTTATTGCT
AAAAATCCATTTATACTTCCCCTTATCAATCAACGCTTAGAACGAAAAGAA
ATTTGGCGTGAATATTGGGCTTTAGTTGAAGGAAAATTTTCACCTAAGCAT
CAAGTTTTGAGAGACAAAATTGGACGGAACCGTCATGACAGACGTAAACG
AATCATTGATTCTAAAAACGGTCAACATGCTATGACAATCATTGACGTTTT
20 GAAGTATATCCAAAATAGTAGTCTCATAAAATGCCGACTGGAAACCGGAA
GAACCCATCAAATTCGCATTCACCTTATCTCATCACGGACATCCTTTAATAG
GAGATCCCCTCTACAACCCTTCTTCTAATAATGAAAGGTTAATGCTACACG
CTCACCGATTGACTCTATCCCATCCATTAACCTTGCGAAACTATTAGCGTAG
AGGCCCTTCATCTACTTTCGAGAAGGTTTTAAACAATTATAAAAAAAGGAG
25 TTGGATAA

(SEQ ID NO: 62)

MLVSSLVSCSFFLVISSLSSSTRNKSINLSNNFIDLLYQFSIKCYNKTMSFSLKIR
NPYGEHTVKELLEDYFLIPRKIRHFLRVKKHVLINNEFINWQTVVQENDTITLIF
DDEDYPTKKIPLGRAELIDCLYEDEHLIIVNKPEGMKTHGNQPNEIALLNHVSA
30 YSGQTCYVVHRLDMETSGAVLFAKNPFILPLINQRLERKEIWREYWALVEGKF
SPKHQVLRDKIGRNRHRRRKRIIDSKNGQHAMTIIDVLKYIQNSSLIKCRLETG
RTHQIRIHLSHHGHPLIGDPLYNPSSNNERLMLHAHRLTL SHPLTCETISVEAPS
STFEKVLNNYKKGVG*

35

Sequence description:

- 40
- A] Length: 1029 bp - 343 aa (Full length gene sequence)
 - B] No obvious Shine-Dalgarno sequence upstream of the putative TTG start codon. Possesses a potential leader peptide sequence.

ID-107

Clone 2-54

5

(SEQ ID NO: 63)

GAAGTAAATGCAACTCAACCTAATAATAGAACTACCTATATTATACCCGAA
AGCAGTCATTCCATTGCAGAACACAGAGATTCCTGATAGAATCAAAGGG
TTCTTCGGTTGCATTACTTAATAGCGATGAATTTAGAAAGACAGCGGGAGA
10 GGATAGAGGTTTTGAAAGGGATAAGTTGAGGTCTTTGGATATCATTCTAA
GGGAGATTTATCGACAAGTAATGTCATAGGTAATACGGACATTGCTAGTC
AGATATCGTTGGGCTTTAAAAAGAATGCGATGCAGGAACACCATCTTACT
AAAACATTCTCTCAAAAGGATGGAAAGTTATCGTCTGTTATAGAGGGGGAT
GCTTGCTATTGGCAAAGAGAGAAAGTAGAGAAAGAAATAAAATATAGTGGTA
15 ATTTATGGCAAAAATTAAGCTAAGGCACACTGCCTTGTTTGCTGTGTTG
ATAATTTGAATTTTGAAGATATAAAATCTTATTTTCAATATTATTGTCATCT
AAACCATCAGCTCAAATTACCTAAAGGTGCTATACTTTCTGCTAAAACAGA
AGTATATAGGGGAGGAGATTTTGGGAGAAAAAATAAAGATAATGTGTTTG
GTTACCGTATCCCCTCATTATTGAAAACCCAAAAAGGAACCTTACTTGCGG
20 GAGCTGATGAAAGAATTGAGCAAGCTTGTGATTGGGGAAACATAGGAATG
GTTATTCGCCGTAGTGAGGATGATGGTGTCACTTGGGGAAAAAGAGAAAC
TATTGTCAATCTCCGTAATAACCCTAGAGTTCCGCTAGTTACTAGTGGTGA
CTATAGTGGCTCACCTATTAATATGGATATGGCATTAGTTCAAGATACTAG
CTCCAAGACGAAACGTATTTTTTCAATATATGATATGTTTCCAGAAGGAAG
25 AGGCGTTATTAGTATTGCTAACACACCTGAAAAAGAATATACCCAAATCG
GAGGACAGTCTTATCTTAATTTATATAATAATGGAAAGAAATCGAAGGTTT
TACTATCCGTGACAAAGGTATTGTATATAATTTTAAAGGGAAAAAGACTG
ATTATCATGTTATAACAGAACTACTAAAAGTGACCATTCAAATCTAGGGG
ATATTTATAAGGGAAAAACAGCTACTTGGAAATATATATTTTACAAAACATA
30 AACGTCACCATTTCGTTTAGCAAAATCAAGCTATGTGTGGATGTCATATA
GCGATGATGATGGTAGGACATGGTCATCACCTAGAGATATAACAGCAAGT
CTTCGTCAGAAAGGCATGAAATTTTTTGGGAATAGGACCTGGAAAAGGTAT
AGTTTTTAAATGGGGGCCACACGCTGGTCGTATTATTATTCCTGCCTATTCT
ACGAATTGGAAATCTCATCTAAGAGGTTTACAATCTTCACGCCTAATTTAT
35 TCAGACGACCATGGAAAAACGTGGCATACTGGAAAAGCAGTTAATGATAA
CCGTATACTTTCTAATGGTGAAAAAATTCACCTCCTTAACAATGGATAATAA
AAAAGAACAAAATACAGAATCCGTACCCGTTCAATTGAAAAATGGGGACA
TTAAGTTATTTATGAGGAATCTAACTGGTAACCTAGAAGTAGCCACAAGTA
AAGACGGCGGGGAGACTTGGCAAAACCATGTTAAACGATATAAGGAAATT
40 CATGATGCTTACGTCCAACATCAGCTATTCGCTTTGAGCATGACAAAAAA
GAGTATATTTTATTAGTGAATGCTAATGGGGCCAGGGAAGAAGTGCCAAGA
TGGATATGCACGTCTAGCGCAAGTTAATCGAAATGGTAGTTTTAAGTGGTT
ATATCACCATCACATTCAAGATGGTTCGTTTGCTTACAACCTCTGTTCAACA
ACTTAATAATGATCAATTTGGTGTCTTTATGAACATAGAGAAAAACATCA

AAATAGTTTTACTTTAAATTACAAAGTTTTTAATTGGAGTTTTCTTAGTCAA
AATACAGAGAAGCAAGGCACTTTATGGGAGAAAAATGGCAGCAAATTGGCA
TGTTTTGTTTAAATTTTATTTATGA

5

(SEQ ID NO: 64)

ELNATQPNRRTTYIIPESHSIAEQQRFLIESKGSSVALLNSDEFKRTAGEDRGF
ERDKLRSLDIIPKGDLSNSVIGNTDIASQISLGFKKNAMQEHHLTKTFSQKDG
KLSSVIEGMLAIGKEKVEKEIKYSGNLWQKLKAKAHCLVCCVDNLNFEDIKS
YFQYYCHLNHQLKLPKGAILSaktevyrGGDFGRKNKDNVFGYRIPSLLKTQ
10 KGTLLAGADERIEQACDWGNIGMVIRSEDDGVTWVGKRETIVNLRNNPRVPL
VTSGDYSGSPINMDMALVQDTSSKTKRIFSIYDMFPEGRGVISIANTPEKEYTQI
GGQSYLNLYNNGKKS KVFTIRDKGIVYNFKGKKTDYHVITETTKSDHSNLGDI
YKGKQLLGNIYFTKHKTSPFRLAKSSYVWMSYSDDDGRTWSSPRDITASLRQ
KGMKFLGIGPGKGIVLKWGPHAGRIIPAYSTNWKSHLRGSQSSRLIYSDDHG
15 KTWHTGKAVNDNRILSNGEKIHSLTMDNKKEQNTESVPVQLKNGDIKLFMRN
LTGNLEVATSKDGGETWQNHVKRYKEIHDA YVQLSAIRFEHDKKEYILLVNA
NGPGKKCQDGYARLAQVNRNGSFKWLYHHHIQDGSFAYNSVQQLNNDQFG
VLYEHREKHQNSFTLNYKVFNWSFLSQNTEKQGTLWEKMAANWHVLKFYFYL
*

20

Sequence description:

A] Length: 2052 bp - 684 aa (partial gene sequence)

B] N-terminus has yet to be determined

25

ID-108

30

Clone 2-61

(SEQ ID NO: 65)

ATGCCTAAATTAATCGTATCTTTCCTCTGCATTTTTATTATCCCTGACTTGTG
TAAACTCTGTGCAAGCTGAAGAACATAAAGATATTATGCAAATTACCCGA
35 GAAGCCGGATATGATGTTAAAGATATTAATAAACCTAAAGCGTCTATCGTT
ATTGACAATAAAGGTCATATTTTGTGGGAAGATAACGCCGATTTAGAACGT
GATCCCGCTAGCATGTCTAAAATGTTTACTTTATATTTACTATTTGAAGACT
TAGCTAAAGGAAAAACAAACCTCAACACCACAGTGACTGCAACAGAAACA
GACCAAGCCATAAGTAAGATTTATGAAATTAGTAATAACAATATTCATGCT
40 GGGGTTGCTTATCCTATTCGTGAACTGATTACTATGACGGCTGTCCCGTCA
TCTAATGTAGCAACTATTATGATTGCTAACCCTTATCACAAAACAATCCT
GACGCCTTTATTAAACGAATCAATGAAACCGCCAAGAACTCGGTATGAC
AAAAACTCACTTTTATAACCCCAGTGGGGCGGTAGCGAGTGCTTTTAATGG
ACTTTACTCCCCAAAAGAATACGATAACAATGCTACTAACGTTACGACTGC

ACGTGATCTATCAATTTTAACTATCATTTTCCTTAAAAAATACCCTGATATA
CTGAACTATACAAAATATCCTGAAGTCAAGGCCATGGTCGGAACCTCCTTAT
GAAGAAACATTTACAACCTTATACTACTCTACCCCCGGCGCTAAATTTGGA
TTAGAAGGAGTAGATGGCTTAAAAACTGGTTCTAGCCCTAGCGCTGCTTTT
5 AATGCCTTAGTTACAGCTAAACGCCAGAATACTCGCTTGATAACTGTGGTT
TTAGGAGTTGGCGATTGGTCAGACCAAGACGGAGAGTACTATCGTCATCC
GTTTGTCAACGCTCTTGTAGAAAAAGGTTTTAAAGACGCTAAAAAATATTTTC
TTCTAAAACTCCTGTATTAAAAGCCGTTAAACCTAAAAAAGAAGTTACTAA
AACCAAACTAAATCTATTCAAGAACAGCCTCAAACAAAAGAACAGTGGT
10 GGACAAAAACAGATCAATTTATCCAATCACATTTTGTATCTATTTTAATTG
TTCTGGGCACCATCGCTAGCCTTTGTCTTTTAGCTGGGATAGTATTACTTAT
AAAGCGCTCTAGATAA

(SEQ ID NO: 66)
15 MPKLIVSFLCILLSLTCVNSVQAEHKDIMQITREAGYDVKDINKPKASIVIDN
KGHILWEDNADLERDPASMSKMFTLYLLFEDLAKGKTNLNTTVTATETDQAI
SKIYEISNNNIHAGVAYPIRELITMTAVPSSNVATIMIANHLSQNNPDFAFIKRINE
TAKKLGMTKTHFYNPSGAVASAFNGLYSPKEYDNNATNVTTARDLSILTYHF
LKKYPDILNYTKYPEVKAMVGTPYEETFTTYNYSTPGAKFGLEGVDGLKTGS
SPSAAFNALVTAKRQNTRLITVVLGVGDWSDQDGEYYRHPFVNALVEKGFK
20 DAKNISSKTPVLKAVKPKKEVTKTKTKSIQEQPQTKEQWWTKTDQFIQSHFVS
ILIVLGTIASLCLLAGIVLLIKRSR*

25 Sequence description:

A] Length: 1188 bp - 396 aa (full length gene)
B] Shine Dalgarno sequence present upstream of
30 ATG start codon, possesses a potential signal
peptide

35 ID-109

Clone 45

(SEQ ID NO: 67)
40 ATGACTGAAAAATATTATAATTGGGCAACGCTTGGAACCGGCGTTATTGCC
AACGAATTAGCCCAAGCACTGGAAGCACGTGGACAAAAATTATATTCTGT
AGCTAATAGAACTTACGACAAAGGACTTGAATTTGCTAACAAATATGGTA
TCCAAAAAGTTTATGATCACATAGATCAAGTATTTGAAGACCCTGAAGTGG
ATATCATTTATATCTCTACTCCCCACAATACTCACATCTCATTTTTACGAAA

5 GGCTTTAGCAAATGGTAAGCACGTTCTTTGCGAAAAATCTATTACTTTAAA
TAGTACTGAGCTTAAAGAAGCCATAGATTTAGCCGAAACTAACCATGTTGT
CTTAGCTGAAGCCATGACTATTTTTCATATGCCAATTTACCGCCAATTAAA
AACATTAGTTGATAGTGGAATAATAGGACCGTTAAAAATGATTCAAATGA
10 ATTTTCGGAAGTTATAAAGAATATGATATGACTAACCGTTTTTTCAGTCGTG
ACCTAGCAGGCGGTGCTTTGCTGGACATTGGTGTTTATGCACTTTCTTGTAT
TCGCTGGTTTATGTCAGAAGCACCTCACAACATTACCTCTCAAGTTACATT
TGCACCAACAGGGGGTTGATGAACAAGTTGGTATCCTACTAACCAACCCAG
CAAATGAGATGGCGACTGTCAGCCTTAGTTTACATGCAAAACAACCTAAA
15 CGAGCAACTATCGCTTACGATAAAGGCTACATTGAACTTTTTGAATATCCG
CGAGGACAAAAGGCAGTTATTACTTATACTGAGGATGGGCATCAAGATAT
TATCGAAGCTGGCAAAACTGAAAATGCTCTCCAATATGAGGTAGCTGATA
TGGAAGAAGCCATTTTCAGGAAAAACTAACCACATGTACTTAAACTATACC
AAAGATGTTATGGATATCATGACACAGCTACGTCAAGAATGGGGATTAC
15 CTACCCAGAAGAAGAAAAATGA

(SEQ ID NO: 68)
MTEKYYNWATLGTGVIANELAQALEARGQKLYSVANRTYDKGLEFANKYGI
QKVYDHIDQVFEDPEVDIIYISTPHNTHISFLRKALANGKHLCEKSITLNSTEL
20 KEAIDLAETNHVVLAEAMTIFHMPYRQLKTLVDSGKLGPLKMIQMNFSGYK
EYDMTNRFFSRDLAGGALLDIGVYALSCIRWFMSEAPHNITSQVTFAPTGVDE
QVGILLTNPANEMATVSLSLHAKQPKRATIAYDKGYIELFEYPRGQKAVITYT
EDGHQDIIEAGKTENALQYEVADMEEAISGKTNHMYLNYTKDVM DIMTQLR
QEWGFTYPEEEK*

25 Sequence description:

30 A] Length: 984 bp - 328 aa (full length gene)
B] Shine Dalgarno sequence present upstream of
ATG start codon, possesses a potential signal
peptide

35 ID-110

Clone 2-2

40 (SEQ ID NO: 69)
GTGTATTCTCCTGTAAATCTTCTAAAGGAAAAGTGATATTGTTAAAAAGT
GATTTTCTAAAGAGCTTCATAGAAAGGAGAGGAAATATTTGTTTT
(SEQ ID NO: 70)
MYSPVKSSKGKVILLKSDFLKSFIERRGNICF

Sequence description:

- 5 A] Length: 96 bp - 32 aa (partial sequence)
 B] GTG start codon - no obvious Shine-Dalgarno
 sequence
 Possesses a potential signal peptide

10

ID-111

15 Clone 2-3

(SEQ ID NO: 71)
AAATACTGTATCATTGCAACCTCAAATGCAGGTTTTGGAAACGAAGCATT
ACAGGTGACAGCGATAAAGACTTGAAAATTATGGAACGAATTTCTCCATA
20 TTTCCGTCCAGAATTTCTAAATCGTTTCAATGGTGTTATTGAATTCTCTCAC
CTAAGCAAAGATGACTTAAGCGAAATTGTAGATTTGATGCTTGATGAAGTT
AACCAAACAATTGGCAAAAAAGGAATTGACCTTGTGGTAGATGAAAATGT
TAAATCACACTTAATTGAACTGGGTTATGACGAAGCAATGGGAGTACGTC
CATTGCGCCGTGTCATCGAGCAAGAAATTCGAGATCGCATCACAGACTACT
25 ATCTCGATCATAACAGACGTTAAACACCTAAAAGCTAATTTGCAAGATGGCC
AAATCGTCATTTCTGAAAGATAA
(SEQ ID NO: 72)
KYCIATSNAGFGNEAFTGDSDKDLKIMERISPYFRPEFLNRFNGVIEFSHL
DLSEIVDLMLDEVNQTIGKKGIDLVVDENVKSHLIELGYDEAMGVRPLRRV
30 QEIRDRTDYYLDHTDVKHLKANLQDGQIVISER*

Sequence description:

- 35 A] Length: 429 bp - 143 aa (partial sequence)
 B] N-terminus yet to be elucidated. This gene
 was not in frame with nuc

40

ID-112

Clone 2-5

(SEQ ID NO: 73)

ATGTCAATGAATTTTTCATTTTACCACAATATTGGTCCTATTTTAATTATG
GTGTGATGGTAACCATTATGATTTCAACATGTGTTGTTTTTTTGGAACTAT
TATAGGCGTGTTAATTGCTTTAGTAAAGCGTACTAATTTACATTTTCTCACA
ATATTAGCTAATTTCTATGTATGGGTATTTTCGTGGGACACCGATGGTAGTT
5 CAAATTATGATTGCTTTCGCATGGATGCATTTTAACAATTTACCAACAATT
AGCTTTGGTGTTTTAGATTTAGATTTTACACGACTTTTACCTGGTATCATT
TCATTTCTTAATAAGTGGTGCCTATATTTTCGGAAATTGTACGTGCAGGGA
TTGAGGCTGTACCATCTGGACAAATAGAAGCAGCTTACTCGTTGGGGATTC
GACCTAAAAATACACTTCGCTATGTTATCTTACCCCAAGCTTTTAAAAATA
10 TTTTACCTGCTCTAGGGAATGAATTTATTACAATTATTAAAGATAGTGCTCT
CCTTCAAACCTATTGGTGTCATGGAATTATGGAACGGAGCACAAATCAGTTGT
AACGGCTACTTACTCACCAGTTGCACCGTTATTATTTGCAGCATTTTACTAT
TTAATGTTGACAACGATTCTCTCAGCTTTGTTAAAACAAATGGAGAAATAT
CTTGGGAAAGGGGTAAAAATAGATGGTTGA

(SEQ ID NO: 74)

MSMNFSFLPQYWSYFNYGVMVTIMISTCVVFFGTIIGVLIALVKRTNLHFLTIL
ANFYVWVFRGTPMVVQIMIAFAWMHFNNLPTISFGVLDLDFTRLLPGIIISLNS
GAYISEIVRAGIEAVPSGQIEAAYS LGIRPKNTLRYVILPQAFKNILPALGNEFITI
IKDSALLQTIGVMELWNGAQSVVTATYSPVAPLLFAAFYYLMLTTILSALLKQ
20 MEKYLKGKGVKIDG*

Sequence description:

25 A] Length: 699 bp - 233 aa (full length gene)
B] Shine-Dalgarno sequence preceded the 'ATG'
start codon. Possesses a potential leader peptide
sequence.

30

ID-113

35 Clone 2-7

(SEQ ID NO: 75)

ATGAAAGACCTATTACGAAATAGTCTAGAGCAAAGTGGAAATTTAAGTTT
TCAAGATATGATTTTACATATTCTTGTAGCAGCTTTATTGAGTGTTAGTTATT
40 TATGTTTCCTATGCTTATACGCATAGTGGAACCTGCCTATAGTAAAAAGTTT
AATGTTTCATTAATGACATTGACGGTCTTGACTGCAACAGTAATGACCGTT
ATTGGTAATAATGTAGCCTTGTCATTGGGTATGGTCGGTGCCTTGTCAGTT
GTTTCGTTTTAGGACAGCCATAAAAGATTCAAGAGATACAGTTTATATTTT
TGGACCATAGTTGTTGGTATCTGTTGTGGTGTGGTGACTATGTGGTAGCT

GCATTAGGAAGTAGCGTTATCTTTATCTTATTATGGGTATGGGACGTGTT
AAAAACGAGAATCGTATGTTATTGATTGTGAAGTGCGATAGAACACTAGA
AGTTGATTTAGAAGGAATTTTCTTCCAATATTTTGACGGAAAAGCTGTTCA
GCGTGTTAAAAATTCAACAATACTATTGAAATGATTTTCGAAATCTC
5 TAGAAAAGATTACGATAAGCAACTCCATGTAGATAATCAGTTAACTGAAA
AAGTGTACCAATTGGGAAATATTGATTATTTCAACATTGTTAGCCAAAGCG
ACGAAATCAATGGGTAG

(SEQ ID NO: 76)
MKDLLRNSLEQSGNLSFQDMILHILVAALLSVVIYVSYAYTHSGTAYSKKFNV
10 SLMTLTVLTATVMTVIGNNVALSLGMVGALSVVRFRTAIKDSRDTVYIFWTIV
VGICCGVGDYVVAALGSSVIFILLWVMGRVKNENRMLLIVKCDRTLVDLEGI
FFQYFDGKAVQRVKNSTTNTIEMIFEISRKDYDKQLHVDNQLTEKVYQLGNID
YFNIVSQSDEING*

15
Sequence description:

A] Length: 678 bp - 226 aa (full-length gene)
B] ATG start codon is preceded by a Shine-
20 Dalgarno sequence-Possesses a potential leader
peptide sequence

25 ID-114

Clone 2-8

(SEQ ID NO: 77)
30 AAAAATTCATTTTAGATTTCATTTTACGACTATATACTCAGAAGTACCAAAC
CTAATCCAAGGTTTGAAAAAAGAAAGAAGGAAGTCAGTATGACAAACTAT
AAAAACAACTTTAAAGATGAGGCTATACGTGTTGAAGAGACAACAAAAGA
ATCATTTTACGATGTTGATATTGCCTTGTTTTTCAGCTGGTGGATCTATTTCA
GCAAAGTTCGCTCCTTATGCAGTAAAGTCTGGAGCAGTTGTAGTAGATAAC
35 ACGTCATATTTTCGTCAGAATCCTGATGTTCCACTAGTTGTTCCCTGAAGTAA
ATGCTCATGCCATGATTGGTCATAATGGTATCATAGCTTGTCCCAATTGTTC
TACTATTCAAATGATGATTGCTTTAGAGCCCATTCGTCAAAAATGGGGGAT
AGAGCGTGTTATAGTTTCCACCTATCAAGCTGTTTCGGGTTCAGGTGCACG
TGCTGTTGAAGAACTAAGGAACAGTTGAGACAAGTTT

40 (SEQ ID NO: 78)
KFILDSFYDYILRSTKPNPRFEKRKKEVSMTNYKNNFKDEAIRVEETTKESFYD
VDIALFSAGGSISAKFAPYAVKSGAVVVDNTSYFRQNPDVPLVVPEVNAHAMI
GHNGIACPN CSTIQMMIALEPIRQKWGIERVIVSTYQAVSGSGARAVEETKEQ
LRQV

Sequence description:

- 5 A] Length: 499 bp - 165 aa (partial sequence)
 B] N-terminus has yet to be determined

ID-115

10

Clone 2-9

15 (SEQ ID NO: 79)
 ATGACAAATGAATTGATAATGCAAGCTTTTGAGTGGTATTTACCTAGTGAT
 GGGAATCACTGGAAGAAATTAGAGGAGTCTATATCAGACCTTAAAAAACT
 TGGAATTAGTAAAATCTGGTTACCACCAGCATTTAAGGGAACTAGCAGTG
 ATGATGTAGGATATGGTGTTTATGATCTCTTTGATTTAGGAGAATTTGACC
 AGAATGGAACAATTAGAACAAAATATGGTAGGAAAGAAGAGTATCTAAA
20 GCTTATTAAGTCGTTAAAGGCAAATGGCATTAAACCGTTTGCAGATATCGT
 TCTTAACCATAAAGCCAATGGTGATCATAAAGAAAAAATTTCAAGTCATCA
 AAGTCAATCCTGAAAATCGTCAAGAAGCATTAAAGTGAACCCTATGAGATT
 GAAGGATGGACGGGATTTGATTTCACAGGTAGACAGGGTGAGTACAATGA
 TTTT

25 (SEQ ID NO: 80)
 MTNELIMQAFEWYLPDGNHWKKLEESISDLKKLGISKIWLPPAFKGTSSDDV
 GYGVYDLFDLGEFDQNGTIRTKYGRKEEYLKLIKSLKANGIKPFADIVLNHKA
 NGDHKEKFQVIKVNPNRQEALSEPYEIEGWTGFDFPGRQGEYNDF

30

Sequence description:

- A] Length: 456 bp - 152 aa (partial sequence)
 B] ATG start codon is preceded by a Shine-
35 Dalgarno sequence, no leader peptide sequence.

ID-116

40

Clone 2-10

(SEQ ID NO: 81)
ATGGAGGTTCTTATGAAGAAAGTGTTAGTAAGTAGTCTTTTGGTTTTAGGG
ATTACGATAACGTTACAACCAGTAGTTGAGGCTAAGGGGCCAAAAGTAGC
TTATACACAAGAGGGAATGACTGCTCTTTCGGACACAAATAAAGATAAAG
TCACTACTATTTCTATTGACGAGATTCAAAAAAGCTTAGAAGGTAAGAAGC
5 CGATTACTGTTAGTTTTGATATTGATGATACACTGCTTTTCAGTAGTCAATA
TTTTCAATATGGTAAAGAATATGTAACCTCCTGGATCGTTTGATTTTCTTCAT
AAACAAAAATTCTGGGATCTTGTTGCAAAACGAGGAGATCAAGATTCCAT
TCCCAAAGAATATGCTAAAAAATTAATTGCTATGCATCAAAAACGAGGAG
ATAAAATTGTTTTTATAACAGGTAGGACAAGAGGGTCAATGTATAAGGAG
10 GGCGAGGTTGATAAAACAGCTAAAGCCTTAGCTAAAGATTTTAAATTTGTA
CCATCTGAT

(SEQ ID NO: 82)
MEVLMKKVLVSSLLVLGITITLQPVVEAKGPKVAYTQEGMTALSDTNKDKVT
TISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFQYGKEYVTPGSFDLHKQKFW
15 DLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGSMYKEGEVDKTA
KALAKDFKFVPSD

20 Sequence description:
A] Length: 516 bp - 172 aa (partial sequence)
B] ATG start codon is preceded by a Shine-
Dalgarno sequence, Possesses a leader peptide
sequence.

25 ID-117

30 Clone 2-17

(SEQ ID NO: 83)
ATGCTTAAAAGATTATTTACTGAAGATGGGGAATTGACAAAGATTAGTCGT
CGTTTCGTTTGGATGTTAGTGGTTATCTATTGTCTTATTATTGTCAGGATGT
35 GTTTTGGGCCTCAAATTATGATTGAGGGGGTATCAACTCCGAATGTTTCAGC
GCTTCGGAAGAATTGTAGCTCTTTTAGTACCATTTAATTCTTTTCGTAGTTT
AGATCAGCTAACTAGCTTTAAAGAGATTCTTTGGGGTTATTGGTCAAAATGT
AGTGAATATTTTACTGCTGTTTCCTCTCATTATAGGGTTACTATCCCTAAAG
CCAAGTTTACGGAAATATAAAAGCGTTATATTACTTGETTTCTTGATGTCTC
40 TTTTCATAGAGTGTACTCAAGTTGTTTTAGATATTTTAATAGATGCTAATCG
GGTTTTTGAAATCGACGATCTATGGACAAATACCTTAGGCGGTCCTTTTCGC
CCTATGGAGTTATCGAAACATAAAAGGTTGGCTTCTAACTATTAGAAAATG
A

(SEQ ID NO: 84)

MLKRLFTEDGELTKISRRFVWMLVVIYCLIIVRMCFGPQIMIEGVSTPNVQRFG
RIVALLVPFNSFRSLDQLTSFKEILWVIGQNVVNILLFPLIIGLLSLKPSLRKYK
SVILLAFMLSLFIECTQVVLDILIDANRVFEIDDLWTNTLGGPFALWSYRNIKG
WLLTIRK*

5

Sequence description:

- 10 A] Length: 516 bp - 172 aa (full-length gene)
 B] ATG start codon is preceded by an Shine-
 Dalgarno sequence. Possesses a potential leader
 peptide sequence. C-terminus need further
 confirmation.

15

ID-118

20

Clone 3-3

(SEQ ID NO: 85)

ATGAAAAAGCTTACTTTTATTTGGGATTTAGATGGGACATTAATAGATTTCG
TATGTACCAATTATGGAAGCTCTTGAAGAAACCTATCGTCATTTTGGCTTA
ATATTTGATAAAGAATTAATCCATGAATATATTTTACAGGAATCAGTGGGG
25 CAATTATTGGTAAACCTTTCAGAGGAAGAGCAAATACCTCATGAAAAACT
GAAAGCATATTTTACAAAAGAACAAGAAAGTCGAGATTCTAAAATACATT
TAATGCCATATGCAAAAAGAGATTTTAGAATGGACCAAAGAACAAGATATT
CCCAATTTTATGTATACACATAAAGGAGCAAGTACGCATTCAGTGTTGGAA
ACCTTGCAGATCTCTCATTATTTTGATGAAATTTTAACTGGTGTTCGGGGAT
30 TCGAGCGAAAACCACATCCACAAGGGATTAATTATTTAGTTAAACGATATT
CTTTAGATAAATCAATGACTTATTACATAGGAGATCGTCCACTAGATTTGG
AGGTTGCTCAAAATGCTGGTATAAAATCCATAAACTTAAGGTTAGAGAATT
CCAAAGAAAACCTATAATATTTCAAGTCTCAAAGATATAATATCACTTGATT
TCACTCGTTTGGATTAA

35

(SEQ ID NO: 86)

MKKLTFIWDLDGTLIDSYPIMEALEETYRHFGLIFDKELIHEYILQESVGQLL
VNLSEEEQIPHEKLKAYFTKEQESRDSKIHLMPYAKEILEWTKEQDIPNFMYTH
KGASTHSVLETQISHYFDEILTGVSGFERKPHPQGGINYLVKRYSLDKSMTYYI
GDRPLDLEVAQNAGIKSINLRLNSKENYNISSLKDIISLDFTRLD*

40

Sequence description:

- A] Length: 627 bp - 209 aa (Possible Full-length gene)

B] ATG start codon is preceded by an possible Shine-Dalgarno sequence. No obvious leader peptide sequence.

5

ID-119

Clone 3-7

10

(SEQ ID NO: 87)

ATGGAAAAAGAAAAAAATTAGGTCTTTTACCACTAACAATGCTTGTCATT
GGCTCTCTTATCGGTGGCGGAATCTTTGATTTAATGCAAAATATGAGTTCC
AGAGCCGGTTTGGTACCAATGCTTATTGCTTGGGTAATTACTGCTATCGGG
15 ATGGGAACCTTTCGTTTTAAGTTTTCAAAATTTATCTGAAAAAAGGCCGGAC
CTAACAGCTGGAATCTTTAGTTACGCTAAAGAGGGGGTTTGGAACTTTATG
GGATTAACTCTGCATGGGGTTATTGGTTATCAGCTTGGCTTGGAAATGTT
GCCTACGCTGCACTCTTATTCAGTTCACCTCGGTTATTTCTTTAAATTCTTTG
GTAATGGAAATAATATCATCTCAATTATTGGAGCAAGTATAGTTATTTGGG
20 TTGTCCATTTCTTAATTTTAAGAGGTGTTAATACAGCTGCATTTATTAATAC
CGTAGTTACCTTTGCAAAATTAGTACCTGTTATTATTTTCTTAATTTTCAGCG
TTATTAGCTTTCAAATTTAACATTTTATAGTCTTGATATCTGGGGGAAATGGAT
TACATCAATCAATTTTCAACCAAGTCAATTCAACTATGAAAACCGCTGTTT
GGGTATTTATTGGTATTGAGGGCGCCGTTGTCTTCTCAGGTCGTGCTAAAA
25 AACACTCTGATATTGGTAAAGCAAGTATCCTAGCATTATTCACTATGATT
CACTTTATGTATTGATTTCTGTTTTATCACTTGGTATCATGTACGTCCAGA
ACTTGCAAACCTTAAAAACACCAGCTATGGCTTACGTTCTAGAAAAAGCTGT
TGGTCACTGGGGTGCTATCTTAGTTAACCTTGGTGTTATCATTTCAGTATTT
GGCGCTATTCTTGCTTGGACTTTATTTGCAGCAGAATTACCATATCAAGCT
30 GCTAAAGAAGGTGCTTTTCCTAAATTTTTTGCAAAAGAAAATAAAAACAA
AGCTCCAATCAACTCACTCTTAGTCACTAATCTTTGTGTACAAGCATTCTTA
ATCACGTTCTTATTCACACAAAGTGCTTATCGTTTTGGTTTTCGCATTAGCAT
CATCTGCTATCTTAATTCCTTATGCTTTTACAGCACTATATCAATTACAATT
CACACTCCGTGAGGATAAGTCAACTCCAGGACATCAAAAGAATTTAATTA
35 TCGGTATCCTCGCTACAATCTATGCTGTTTACCTTATCTACGCTGGTGGTTT
TGATTACTTACTTTTGACAATGATTGCTTATACTCTAGGTATGATTCTCTAT
ATTAAAATGAGAAAAGATGACAAGCTTGGCGTAATCATGGTCATAGCTGT
TTCCAGTGTGAAATTGTTATCC

40

(SEQ ID NO: 88)

MEKEKKLGLLPLTMLVIGSLIGGGIFDLMQNMSSRAGLVPMLIAWVITAIGMG
TFVLSFQNLSEKRPDLTAGIFS YAKEGFGNFMGFNSAWGYWLSAWLGNVAY
AALLFSSLGYFFKFFGNGNNIISIIGASIVIWVVHFLILRGVNTAAFINTVVTF
LVPVIIFLISALLAFKFNIFSLDIWGNGLHQSFNQVNSTMKTAVWVFIGIEGAV

5 VFSGRAKKHSDIGKASILALFTMISLYVLISVLSLGIMSRPELANLKTPAMAYV
LEKAVGHWGAILVNLGVIIISVFGAILAWTLFAAELPYQAAKEGAFPKFFAKEN
KNKAPINSLLVTNLCVQAFLITFLFTQSA YRFGFALASSAILIPYAFTALYQLQF
TLREDKSTPGHQKNLIIGILATIIYAVYLIYAGGFDYLLLTMIAYTLGMILYIKMR
KDDKLGVMVIAVSSVKLLS

10 Sequence description:

A] Length: 1356 bp - 452 aa (partial sequence)
B] ATG start codon is preceded by an possible
Shine-Dalgarno sequence. Possesses a potential
leader peptide sequence.

15

ID-120

20

Clone 3-8

(SEQ ID NO: 89)
25 ATGAAATTTGAAAAACGGCAGGTCTATTATGTTGTCATAACATTTGCTATT
TGCTATGCTATACAGGCTTATTGGGGAGCTGTTTCTAATATTTTAACTACGC
TTCATAAGGCAATATTTCCCTTTTTTGATGGGAGCTGGAATTGCCTATATTAT
TAATATTGTAATGTCAGTCTATGAGCGATTATATATAAAGCTTTTTTAAAGG
ATCTAGACTATTAATGGCAATCAAGCGTAGTGTTTCTATGATTTTATCCTAT
GCAACTTTTATTGGTTTAATTGTCTGGCTATTTTCAATTGTCATTCCAGATT
30 TGATTTCTAGTTTGAGTTCTTTATTGGTTATTGATACCGGAGCACTTGCTAA
ATTGGTTAATAATCTCAATGAAAATAAACAATTTCTGAGGCTTTAAATTA
TATGGGAACAGATAAAGACTTAGTTTCTACTTTAAGTGGTTATAGCCAGCA
GATTTTGAAGCAAGTTTTATCTGTTTTAACAATTTACTAACCTCAGTTTCC
TCTATTGCGGCAACACTTCTGAATGTTTTTGTAGTTTTATTTTTTCAATTTA
35 CGTTTTGGCAAACAAGGAGCAGTTGGGACGTCAATTTAATTTGTTAATTGA
TACCTATTTAGGTTCAACAGGCAAAACATTCCATTACGTTTCGTCATATCCTT
CATCAACGTTTCCATGGTTTTTTTTGTAAGCCAACTTTAGAAGCTATGATTT
TAGGAAGTTTGACGGTTATTGGTATGTTGATCTTCCAATTTCCCTTATGCTTT
AACAGTTGGGGTTTGTAGTTGCTTTTACAGCTCTAATACCGGTTGTGGGAGC
40 CTACATTGGTGTTACAATCGGTTTCATCTTAATTGCTACTGAATCGCTTACT
GAAGCATTCTTGTTTGTCTTTTCTTGATCCTTTTACAACAATTTGAGGGGAA
ATGTCATTTATCCGAAAGTTGTCGGTGGATCGATTGGACTGCCTTCTATGT
GGGTTTTAATGGCTATTACTATCGGAGGTGCTTTATGGGGGATCTTAGGCA

TGTTACTTGCTGTTTCCTGTTGCAGCTACTATCTATCAGATTGTAAAAGATCA
TATTATCAAGCGACAAACGCTTAGAAATCGTGCACGAACCTATCGTTAA

(SEQ ID NO: 90)

5 MKFEKRQVYYVVITFAICYAIQAYWGAVSNILTLHKAIFPFLMGAGIAYINI
VMSVYERLYIKLFKGSRLMAIKRSVSMILSYATFIGLIVWLFSIVIPDLISLSS
LLVIDTGALAKLVNNLNENKQISEALNYMGTDKDLVSTLSGYSQQILKQVLSV
LTNLLTSVSSIAATLLNVFVSFIFSIYVLANKQLGRQFNLLIDTYLGSTGKTFH
YVRHILHQRFHGFFVSQTLEAMILGSLTVIGMLIFQFPYALTVGVLVAFALIP
VVGAYIGVTIGFILATESLTEAFLFVLFLILLQQFEGNVIYPKVVGGSIGLPSM
10 WVLMATIGGALWGILGMLLAVPVAATIIYQIVKDHIKRQTLRNRARTYR*

Sequence description:

15 A] Length: 1134 bp - 378 aa (full-length gene)
B] ATG start codon is preceded by an typical
Shine-Dalgarno sequence. Possesses a potential
leader peptide sequence.

20

ID-121

Identical to ID-68, as described in WO 00/06736

25

ID-122

30

Clone 3-16

(SEQ ID NO: 91)

35 GTGATTACAATTAAAAAGGAATCTGTTATCAAACCTATTGAAGTATGCTTTT
GGCATTATAATGGGATTTATTATCTTAGCTATTGTAATAGGTGGGCTCCTA
TTTGCATACTACGTTAGTCGTTCTCCGAAATTAACCGATCAAGCTTTAAAA
TCCGTAACTCTAGTTTGGTTTATGATGGTAATAATAAACTTATTGCCGATT
TAGGCTCAGAAAAGCGTGAAAGTGTTAGTGCGGATAGCATTCCACTAAAT
TTGGTTAACGCTATCACTTCTATAGAAGATAAACGTTTCTTTAAACATAGA
40 GGTGTCGATATTTATCGTATTTTAGGTGCAGCTTGGCATAACCTTGTTAGTA
GTAATACGCAAGGTGGTTCAACCCTTGATCAACAGTTGATTAAACTGGCTT
ACTTTTCTACCAATAAATCTGACCAAACGTTAAAACGTAAATCACAGGAA
GTTTGGCTTGCGCTTCAAATGGAGCGTAAATACACCAAAGAAGAAATTCTT
ACTTTCTATATTAATAAAAGTTTATATGGGAAATGGGAATTATGGTATGAGA

ACAACAGCTAAATCATACTTTGGTAAAGACCTAAAGGAATTATCTATTGCA
CAACTTGCTTTGCTCGCTGGTATTCCTCAAGCACCTACACAATATGACCCTT
ATAAAAACCCAGAATCTGCTCAAACAAGACGTAATACCGTTCTTCAGCAG
ATGTATCAAGATAAAAAACATTTCTAAAAAGGAATACGACCAAGCTGTTGC
5 AACTCCAGTAACTGATGGCTTAAAAGAATTAAAGCAAAAATCTACTTATCC
AAAATATATGGATAACTACTTAAAACAAGTTATTAGTGAAGTTAAACAAA
AAACTGGTAAAGATATCTTTACTGCTGGGCTAAAAGTGTATACTAATATCA
ACACTGATGCACAAAAACAACCTATATGACATCTACAACAGTGATACTTAC
ATCGCTTATCCAAACAATGAATTACAAATAGCATCTACCATCATGGATGCG
10 ACTAATGGTAAAGTCATTGCACAATTAGGCGGGGCGTCATCAGAATGAAAA
TATTTTCAATTTGGGACAAATCAATCTGTCTTAACAGACCGCGATTGGGGTTC
TACAATGAAACCTATCTCAGCTTATGCACCTGCTATTGATAGTGGTGTCTA
TAATTCAACAGGTCAATCATTAAACGACTCAGTTTACTACTGGCCTGGTAC
TTCTACTCAACTATATGACTGGGATCGTCAATATATGGGTGGATGAGTAT
15 GCAGACCGCTATTCAACAATCACGTAACGTCCCTGCTGTCAGAGCACTTGA
AGCCGCTGGATTAGACGAAGCAAAAATCTTTCCTTGAAAAATTAGGCATAT
ACTATCCAGAAATG

(SEQ ID NO: 92)
MITIKKESVIKLLKYAFGIIMGFILAIIVIGLLFAYYVSRSPKLTDQALKSVNSS
20 LVYDGNNKLIADLGSEKRESVSADSIPLNLVNAITSIEDKRFFKHRGVDIYRILG
AAWHNLVSSNTQGGSTLDQQLIKLAYFSTNKSDQTLKRKSQEVWLALQMER
KYTKEEILTFYINKVYMGNGNYGMRTTAKSYFGKDLKELSIAQLALLAGIPQA
PTQYDPYKNPESAQTRRNTVLQQMYQDKNISKKEYDQAVATPVTDGLKELK
QKSTYPKYMDNYLKQVISEVKQKTGKDIFTAGLKVYTNINTDAQKQLYDIYN
25 SDTYIAYPNNELQIASTIMDATNGKVIAQLGGRHQNENISFGTNQSVLTDRDW
GSTMKPISAYAPAIIDSGVYNSTGQSLNDSVYYWPGTSTQLYDWDRQYMGWM
SMQTAIQSRNVPVRALEAAGLDEAKSFLEKLGIIYPEM

30 Sequence description:

A] Length: 1386 bp - 462 aa (partial sequence)
B] GTG start codon is preceded by an
typical Shine-Dalgarno sequence. Possesses a
35 potential leader peptide sequence.

ID-123
40

Clone 3-17

(SEQ ID NO: 93)
ATGGCTAATGTATATGATTTAGCAAATGAATTAGAACGTGCTGTTTCGTGCT
TTACCAGAATACCAAGCAGTTTTTAAGTCAAAAAGCAGCTATTGAAAATGA
TGCGGATGCACAAGTGCTTTGGCAAGACTTTTTGGCTACCCAATCAAAAGT
TCAAGAAATGATGCAATCTGGCCAAATGCCAAGTCAAGAAGAACAAGATG
5 AAATGTCTAAACTTGGGGAAAAAATTGAATCCAATGACCTTTTAAAAGTTT
ATTTTGACCAACAACAACGGTTGTCTGTCTATATGTCTGATATCGAAAAAA
TTGTCTTTGCACCCATGCAGGACTTGATGTAA

(SEQ ID NO: 94)
10 MANVYDLANELERAVRALPEYQAVLTAKAAIENDADAQVLWQDFLATQSK
VQEMMQSGQMPSQEEQDEMSKLGEKIESNDLLKVYFDQQQLSVYMSDIEKI
VFAPMQDLM*

Sequence description:
15 A] Length: 336 bp - 112 aa (full length sequence)
B] ATG start codon is preceded by an
typical Shine-Dalgarno sequence. No obvious
potential leader peptide sequence.
20

ID-124

25 Clone 3-26

(SEQ ID NO: 95)
30 ATGGCAGAAATCACAGCTAAACTTGTAAGAATTGCGTGAAAAATCAGG
TGCAGGCGTTATGGACGCTAAAAAAGCATTAGTAGAACTGATGGTGACC
TTGATAAAGCGATTGAATTACTTCGCGAAAAAAGGTATGGCTAAAGCAGCT
AAAAAAGCAGACCGTGTTGCTGCTGAAGGTTTAACAGGTGTTTATGTTGAT
GGTAACGTTGCAGCAGTTATTGAAGTTAA

(SEQ ID NO: 96)
35 MAEITAKLVKELREKSGAGVMDAKKALVETDGDLDKAIELLREKGMAKAAK
KADRVAEGLTGVYVDGNVAIV

Sequence description:
40 A] Length: 230 bp - 76 aa (partial sequence)
B] ATG start codon is preceded by an
typical Shine-Dalgarno sequence. No obvious
potential leader peptide sequence.

5 ID-125

Clone 3-33

10 (SEQ ID NO: 97)
ATGATAAAAAACCTGTTATTAACAGGTTTTTTATCATTTAATGACGGAAAA
CTGGACACAAATTATTTTTCTTGTATAATTAAATATATTATTTCTTATCAGG
AGGTTATGATGACATTAGAGAAACGATTAA

15 (SEQ ID NO: 98)
MIKNLLLTGFLSFNDGKLDTNFYFSCHIKYIISYQEVMMTLEKRF

Sequence description:

20 A] Length: 134 bp - 44 aa (partial sequence)
B] ATG start codon is preceded by an
typical Shine-Dalgarno sequence. Possible
potential leader peptide sequence.

25 ID-126

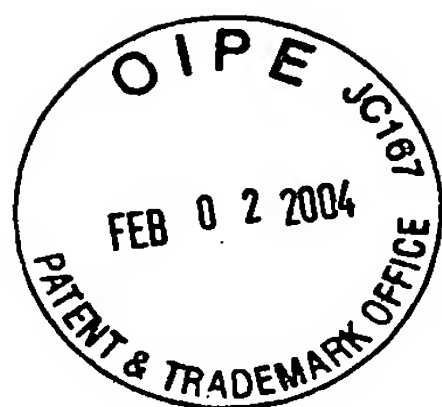
Clone 3-41

30 (SEQ ID NO: 99)
ATGAAAAATAATAAAAAATAATGGTTTTCTGAAAAATTCCTTTATTTACATA
TTATTGATTATTGCGGTTATTACAACCTTTCAATACTATTAA

35 (SEQ ID NO: 100)
MKNNKNNGFLKNSFIYILLIIVITTFQYYL

Sequence description:

40 A] Length: 94 bp - 31 aa (partial sequence)
B] ATG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential
leader peptide sequence.



ID-127

5

Clone 3-42

(SEQ ID NO: 101)
10 ATGTTAGATATTATCTTATCCGGAATTTTCGCAAGGATTACTTTGGTCAATTA
TGGCAATTGGCGTGTTTATCACTTTTCGTATCTTAGACATAGCCGATCTCTC
TGCAGAAGGGGCTTTCCCTATGGGGGCTGCAGTTTGCGCCTTATGTATCGT
TAA

(SEQ ID NO: 102)
15 MLDIILSGISQGLLSIMAIGVFITFRILDIADLSAEGAFPMGAAVCAALCIV

Sequence description:

20 A] Length: 158 bp - 52 aa (partial sequence)
B] ATG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential
leader peptide sequence.

25

ID-128

30 Clone 3-43

(SEQ ID NO: 103)
ATGGAAATGCCTAAAAGAAATGAATTACTCAATAAAGAAATTAAAATGAG
TATTGATAAACTTAGATATAAAGAACCAGAGAGTGAACATGACAAGCGAC
CTACTTTTATTTGGTAGTACTTATACTTGTTACTGTAGCAGTTATATTGTC
35 GTTATTAA

(SEQ ID NO: 104)
MEMPKRNELLNKEIKMSIDKLRYKEPESEHDKRPTFYLVVLILVTAVILSLF

40

Sequence description:

A] Length: 161 bp - 53 aa (full-length gene)
B] ATG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential

leader peptide sequence.

ID-129

5

Clone 3-44

(SEQ ID NO: 105)
10 GTGGTAAGTAAATTGAGTTTAACAACGATTTTTGCATTGCTATTTTCATCA
ATGCTAATTTACGCAACACCTCTTATCTTTACAAGTATTGGGGGAACCTTC
TCTGAACGTGGTGGTATCGTCAACGTTGGTTTAGAAGGAATTATGGTAATT
GGAGCTTTCTCAGGCGTTGTATTAA
(SEQ ID NO: 106)
15 MVSKLSLTTIFALLFSSMLIYATPLIFTSIGGTFSEGGIVNVGLEGIMVIGAFSG
VVF

20 Sequence description:

A] Length: 179 bp - 59 aa (partial sequence)
B] GTG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential
25 leader peptide sequence.

ID-130

30

Clone 3-46/47

(SEQ ID NO: 107)
35 ATGAGAATTATTGCAATAACTGAAAAGGTTATAAAAGAACTGTTTCGTGAT
AAAAGAACACTTGCTATGATGTTTTTAGCACCTATTTTAATTATGTTTTTGA
TGAATGTTATGTTTTCTGCGAATAGTAATACAAAAGTTAAGATTGGAACTA
TTAACGTTAACACGAAGGTCGTTTCAAATTTAGATAATATTAAGCATATTC
AAGTGAGATCATTAAATTTAACTCATCTGCTAAAAAAGCACTCAAATCAA
ATAAAATTGATGCTCTTATTTCGGAGGACAATAAATCTTATACTGTCTTCT
40 ATGCGAATACAGATTCTTCAAAGACGACTTTAACAAGACAAGCTTTTAAA
ACCGCTGTTAATAACAATGAACAGTAAGGAACTGATTTCGCAAGTTAAAATT
TTAGCTAATAAGAATCCGAACTAGCACAATCCTTACAAACTCGCTCCAAA
TATATCAAAGAAAAATATAATTACGGAAATAAAAAATACAGGCTTTTTTGC
AAAAATGATACCAATACTAATGGGATTTATGGTCTTCTTCTTGGTTTTT

(SEQ ID NO: 108)
MRIIAITEKVIKELFRDKRTLAMMFLAPILIMFLMNVMFSA NSNTKVKIGTINV
NTKVVS NLDNIKHQVRSFKFNSSAKKALKSNKIDALISEDNKS YTVFYANTDS
SKTTLTRQAFKTA VNTMNSKELISQVKILANKNP KLAQSLQTRSKYIKEK YNY
5 GNKNTGFFAKMIPILMGFMVFFLVF

Sequence description:
10 A] Length: 558 bp - 186 aa (partial sequence)
B] ATG start codon is preceded by a
possible Shine-Dalgarno sequence. Potential
leader peptide sequence. C-terminus has yet to be
15 determined.

ID-131

20 Clone 3-48

(SEQ ID NO: 109)
25 GTGATTATCGTTATGAGTAAACATCAAGAAATTTTGGAGTACCTAGAAAAT
TTAGCTGTTGGTAAGAGGGT TAGTGTACGCAGTATTTC AAATCATT TAA

(SEQ ID NO: 110)
MIIVMSKHQEILEY LENLAVGKRVS VRSISNHL

30 Sequence description:
A] Length: 100 bp - 33 aa (partial sequence)
B] GTG start codon is not preceded by a
obvious Shine-Dalgarno sequence. No obvious
35 leader peptide sequence.

40 ID-132

Clone 2-c53

(SEQ ID NO: 111)
ATGTATAGAGAAATTACCGCTGTCGAACACGATCGCTTTGTGAGCGAATCC
AACCAAACAAACCTACTTCAATCTCTTAATTGGCCCAAAGTAAAAGACAA
CTGGGGTAGTCAATTACTTGGCTTTTTTTGACGGTGAAACCCAAATTGCCAG
CGCTAGTATTCTCATCAAATCACTTCCTCTTGGCTTCTCCATGCTGTATATT
5 CCGCGTGGACCAATCATGGATTACTCCAATCTAGATATTGTA ACTAAGGTC
CTTAAGGACCTTAAAGCTTTTGGCAAAAAACAAAGAGCTCTCTTTATCAAG
TGTGATCCTCTCATCTATTT

(SEQ ID NO: 112)
10 MYREITAVEHDRFVSESNQTNLLQSLNWPVKDNWGSQLLGFFDGETQIASA
SILIKSLPLGFSMLYIPRGPI MDY SNLDIVTKVLKDLKAFGKKQRALFIKCDPLI
Y

15 Sequence description:

20 A] Length: 326 bp - 108 aa (partial sequence)
B] ATG start codon is preceded by an obvious
Shine-Dalgarno sequence. No obvious leader
peptide sequence.

25 ID-133

Clone 2-c59

30 (SEQ ID NO: 113)
ATGGACAAGAAAAAATCTTAGTAACGGGTATTGTGCCTAAAGAAGGTCT
AAGAAAGCTTATGGACCGATTTGATGTTACTTATTCAGAAGATCGCCCATT
TTCACGTGACTATGTGTTAGAGCATTTATCTGAATATGACGGATGGTTACT
CATGGGACAAAAAGGTGATAAAGAGATGATTGATGCAGGTGAAA ACTTAC
35 AAATTATTTCTTT

(SEQ ID NO: 114)
MDKKKILVTGIVPKEGLRKLMDRFDVTYSEDRPFSRDYVLEHLSEYDGWLLM
GQKGDKEMIDAGENLQIIS

40 Sequence description:

A] Length: 215 bp - 71 aa (partial sequence)

B] ATG start codon is preceded by an obvious Shine-Dalgarno sequence. No obvious leader peptide sequence.

5

ID-134

10

Clone 2-c62

(SEQ ID NO: 115)
ATTTCGAAAGATGACTACCAAAATATTAGTTTTGGACAGGATCCAGAAGTT
GTTGATTATGCTGGTCTGTTTGAAAAACGCCGTCCAGTTTTAGAAAAAGCA
15 GTTAAAAATTTCTTGCAAGAAGAGAGAGCTACGAGAATGCTATCTGATTTC
TTGCAAGAAGAAAAATGGGTAACTGATTTTGCTGAATTTATGGCGATCAA
AGAACATTTTGGTAATAAGGCGCTTCAAGAATGGGATGACAAGGCTATTA
TACGCCGCGAAGAAGAAGCCTTAGCAGGATATCGTCAAAAGCTTAGTGAA
GTGATAAAATATCATGAAGTAACGCAATATTTCTTTTACAAACAATGGTTT
20 GAGTTAAAAGAATATGCTAATGATAAAGGGATTCAAATTATCGGTGATAT
GCCAATCTACGTTTCTGCCGATAGTGTAGAAGTTTGGACAATGCCTGAACT
GTTT

(SEQ ID NO: 116)
ISKDDYQNISFGQDPEVVDYAGLFKRRPVLEKAVKNFLQEERATRMLSDFLQ
25 EEKWVTDFAEFMAIKEHFGNKALQEWDDKAIIRREEEALAGYRQKLSEVIKY
HEVTQYFFYKQWFELKEYANDKGIQIIGDMPIYVSADSVEVWTMPSELF

30

A] Length: 459 bp - 153 aa (partial sequence)
B] More sequencing is required to determine the
N- and C-termini
enzyme). - Streptococcus pneumoniae (63%)

35

ID-135

Identical to ID-108 described in WO 00/06736

40

Clone 2-c63

ID-136

Clone 2-c66

(SEQ ID NO: 117)

5 ATGGCAAAACAGAAAAATAACTGGCGCCGTGTTGGAGTTGGTGTCTTAC
ACTTGCTTCAGTTGCGACTCTTGCTGCATGTGGAAGTAAATCAGCTTCCCA
GGATTCTAATGGAGCGATTAATTGGGCTATTCCAACAGAAATCAATACACT
AGATTTATCTAAAGTTACAGACACTTACTCAAATCTAGCTATTGGTAACTC
TAGTAGTAATTTCTTCGCTTAGATAAAGATGGAAAGACAAGACCAGACTT
GGCTACTAAAGTTGATGTTTCAAAAGATGGCTTAACTTATACAGCTACATT
ACGTAAAGGCTTGAAGTGGTCAGATGGCAGTAACTTACTGCAAAGGATT
10 TTGTTTATTCATGGCAACGTTTAGTTGATCCTAAAACAGCTTCACAATATG
CTTACCTTGCTGTTGAAGGGCATGTGCTTAATGCCGATAAAATCAACGAAG
GACAAGAGAAAGACTTGAATAAGCTAGGTGTTAAGGCAGAAGGCGATGA
CAAAGTTGTTATTACTTTATCTAGTCCGTCTCCGCAATTCATCTACTACCTT
GCATTCACTAACTTCATGCCACAAAAACAAGAAGTTGTTGAAAAATATGG
15 AAAAGATTACGCAACTACTTCAAAAAATACAGTTTACTCAGGACCATATA
CTGTTGAAGGTTGGAATGGTTCGAATGGTACTTTCACGCTGAAGAAAAAC
AAAAATTATTGGGACGCTAAAAATGTAAAAACAAAAGAAGTTCGCATCCA
GACTGTTAAAAAACAGATAACCGCCGTTCAAATGTATAAACGTGGTGAGT
TAGATGCAGCTAATATCTCAAATACTTCTGCTATTTATCAAGCTAATAAAA
20 ATAATAAAGATGTCACAGATGTTCTAGAAGCGACCACTGCCTATATGGAA
TATAATACTACTGGTTCTGTGAAAGGGCTTGATAATGTTAAGATTCGTTCGC
GCCTTAACTTAGCAACTAACCGTAAAGGAGTTGTTCAAGCAGCCGTTGAT
ACAGGCTCAAAACCGGCAATTGCTTTTGCACCTACTGGTTTAGCCAAAACA
CCAGATGGAAGTATTTGGCAAAATATGTTGCCCCAGGTTATGAATATAAT
25 AAAACTGAAGCAGCAAAACTCTTTAGACTA

(SEQ ID NO: 118)

30 MAKQKNNWRRVGVGVLTLASVATLAACGSKSASQDSNGAINWAIPTEINTLD
LSKVTDITYSNLAIGNSSSNFLRLDKDGKTRPDLATKVDVSKDGLTYTATLRKG
LKWSDGSKLTAKDFVYSWQRLVDPKTASQYAYLA VEGHVLNADKINEGQEK
DLNKLGVKAEGDDKVVITLSSPSPQFIYYLAFTNFMPPQKQEVVEKYGKDYAT
TSKNTVYSGPYTVEGWNGSNGTFTLKKKNKNYWDANKVKTKEVRIQTVKKPD
TAVQMYKRGELDAANISNTSAIYQANKNNKDVTDVLEATTAYMEYNTTGSV
KGLDNVKIRRALNLATNRKGVVQA AVDTGSKPAIAFAPTGLAKTPDGTDLAK
35 YVAPGYEYNKTEAAKLFRL

Sequence description:

40 A) Length: 1143 bp - 381 aa (partial sequence)
B) Shine-Dalgarno sequence precedes ATG codon.
Possesses a potential leader peptide sequence.

ID-137

Clone 2-c67

5

(SEQ ID NO: 111)

TTGAGAGTTTATGAAAATAAAGAAGAGTTGAAAAAAGAAATAAGTAAAC
ATTTGAGAAATACATTATGGAATTTAATAA
TATTCCAGAGAATCTAAAAGATAAAAGAATTGATGAAGTTGATAGAACTC
10 CAGCAGAAAACCTTTCTTATCAGGTTGGCT
GGACCAACTTGGTTCTTAAATGGGAAGAAGATGAAAGAAAGGGACTTCAA
GTAAAAACACCATCGGATAAATTT

(SEQ ID NO: 120)

MRVYENKEELKKEISKTFEKYIMEFNNIPENLKDKRIDEVDRTPAENLSYQVG
15 WTNLVLKWEEDERKGLQVKTPSDKF

Sequence description

20

A] Length: 234 bp - 78 aa (partial sequence)
B] TTG start codon is preceded by a
potential Shine-Dalgarno sequence. No obvious
leader peptide sequence.

25

ID-138

30

Clone 2-c70

(SEQ ID NO: 121)

ATGTCAAAGTTTGATAGTCAGAAAATAATTACTCCGATTATGAAGTTTGTC
AATATGCGAGGGATTATTGCACTCAAAGATGGCATGCTAGCAATTTTACCA
35 CTAACAGTTGTTGGGAGTCTCTTTTAAATATTAGGGCAGCTTCCATTT

(SEQ ID NO: 122)

MSKFDSQKIITPIMKFVNMRGIIALKDGMLAILPLTVVGSFLILGQLPF

40

Sequence description

A] Length: 150 bp - 50 aa (partial sequence)
B] ATG start codon is preceded by a potential
Shine-Dalgarno sequence. Possesses a potential

leader peptide sequence.

ID-139

5

Clone 2-c71

(SEQ ID NO: 123)
10 GAGACCACTTCATCAGTTAAACCAGCAGGAATTGACCGTATCAATCATACC
TCAACACCCCCGAAGAAAACACCCCCAACATTGCAACGACGCATAGCTT
CAAAGATCGTTGTGATACTTTAGAAAGAATTCACAATGAAGACATTGATGT
TTGTTCTGGATTCAATTTGTGGTATGGGAGAGAGCGATGAGGGGCTCATCAC
ATTAGCTTTCAGACTAAAAGAACTGAACCCCTATTCTATCCCTGTCAATTTT
15 TTAGCTTGCTGTTGAAGGAACACCTCTTGGAAAATATAACTATTTGACTCCC
ATTAAATGCTTAAAAAATTATGGCCATGTTGCGTTTTGTTCCTTTCAAGG
AATTAAGATTAAGTGCTGGACGGGAGGTCCATTTTGAGAATTTTGAATCAT
TAGTCACCTTACTTGTTGACTCAACTTTTTTGGGAAATTACCTAACAGAGG
GGGGTCGCAATCAACATACCGATATTGAATTCTTGGAAAAATTACAATA
20 AATCATACTAAAAAGGAATTAATTT
(SEQ ID NO: 124)
ETTSSVKPAGIDRINHTSTPPKKTTPNIAATTHSFKDRCDTLERIHNEIDVCSGFI
CGMGESDEGLITLAFRLKELNPYSIPVNFLLA VEGTPLGKYNLYLTPIKCLKIMA
MLRFVFPFKELRLSAGREVHFENFESLVTLLVDSTFLGNYLTEGGRNQHTDIEF
25 LEKLQLNHTKKELI

Sequence description:

30

A) Length: 535 bp - 178 aa (partial sequence)
B) N- and C-termini require verification

35

ID-140

Clone 2-c73

40

(SEQ ID NO: 125)
ATGCCGGTTTGGACTGCACAGTCTATTCCAAAGGCATTTTGTAGAAAAGCAT
AATACTAAGGAAGGCACCTGGGCAAACTAACCATTCTAAGTGGTTCTTTA
GTATTTTACCAGTTATCTCCTGATGGAGAGGAAATCTCGCGGCATATTTT

5 GATGCTAGTAGTGATATTCCTTTTGTGATCCACAAGTCTGGCATAAAGTT
TCGCCGAATAGTCCAGACTTAAGTTGCTATCTAACTTTTTACTGCCAAAAA
GAAGATTACTTCCATAAAAAAATATGGTCTCACGCGCACACATTCTGAGGTT
ATCGCCAGTGCACCTCTCTTATCTGAGAAGAGTAATATATTAGACCTTGGG
TGTGGTCAAGGGGCGAAACTCACTTTATTTATCGCTGCTGGGACATCAAGTG
ACTTCTGTCGATTCAAACGGACAGAGCCTTGTAGCTTTAGAAAATATGGCA
TTAGAAGAAGAGCTTCCTTACAATATAAAAAGGTATGATATTAATACTACT
GCTATTGAAGGGCACTATGATTTTATTTTATCAACTGTGGTATTTATGTTTT
T

10 (SEQ ID NO: 126)
MPVWTAQSIPKAFLEKHNTKEGTWAKLTILSGSLVIFYQLSPDGEEISRHIFDAS
SDIPFVDPQVWHKVSPNSPDLSCYLTFYCQKEDYFHKKYGLTRTHSEVIASAP
LLSEKSNILDLGCGQGRNSLYLSLLGHQVTSVDSNGQSLVALENMALEEELPY
NIKRYDINTTAIEGHYDFILSTVVFMF

15

Sequence description:

20 A] Length: 563 bp - 187 aa (partial sequence)
B] N- and C-termini require verification

25 ID-141

Clone 2c76

30 (SEQ ID NO: 127)
ATGACAAAGCAAATAATTGCCATTTGGGCTGAAGATGAAGACCATTTGAT
TGGAGTTAATGGCGGTTTACCATGGAGGCTTCCTAAAGAGTTACATCACTT
CAAAGAAACGACCATGGGGGAGGCTTTGCTTATGGGACGAAAGACCTTTG
ATGGAATGAACCGTCGTGTTTTACCTGGTAGAGAGACAATCATCTTAACAA
35 AAGATGAACAATTCCAAGCAGATGGAGTGACAGTCCTAAATAGTGTTGAA
CAAGTTATAAAATGGTTTCAGGAACATAATAAGACCTTATTTATTGTAGGT
GGTGCAAGTATTTATAAAGCATTTCTGCCTTATTGTGAAGCAATCATAAAA
ACTAAAGTTCATGGAAAATTCAAAGGTGATACCTATTTTCCTGATGTTAAT
CTATCTGAGTTT

40 (SEQ ID NO: 128)
MTKQIIAIWAEDDHLIGVNGGLPWRLPKELHHFKETTMGQALLMGRKTFDG
MNRRVLPGRETIILTKDEQFQADGVTVLNSVEQVIKWFQEHNKTLFIVGGASI
YKAFLPYCEAIKTKVHGKFKGDTYFPDVNLSEF

Sequence description:

- 5 A] Length: 417 bp - 139 aa (partial sequence)
 B] ATG start codon is preceded by a Shine-Dalgarno sequence. No leader peptide sequence

10 ID-142

Clone 2-c78

15 (SEQ ID NO: 129)
 TTGTGGCCAAACTGTGCCCCGCTTATTAATAGCACTTTGTTCACCATTGAA
 GATATCTTAACATCAGGTGCTCATAGCAACCCTATTTTAATGGGGGTTATA
 CTTGGCGGGACAATTGTAGTAGTGGCGACAGCACCACCTTTCTTCTATGGCA
 TTGACAGCTATGCTAGGATTAACCGGAATGCCTATGGCTATAGGAGCCTTG
20 TCTGTCTTTGGTTCGTCATTTATGAATGGTGTACTTTTCCATAAATTA
 TTGGAAGTCGTAAAGATAATATAGCTTTTGCTGTTGAGCCTCTAACTCAAG
 CTGACGTGACTTCAGCTAACCCTATTCCAATCTATGTCACTAATTTTGTTGG
 TGGTGCAGCTTGTGGTATTTTAATTGCCTTGATGAAATTAGTTAATGATACT
 CCTGGAACAGCGACACCAATTGCAGGATTTGCTGTCATGTTTGCCTATAAC
25 CCAATGATAAAAGTACTAATAACCGCTCTAGGTTGTATTATCCTATCTTTA
 CTAGCAGGCTATTTTGGAGGCATTGTTTTT
 (SEQ ID NO: 130)
 MWPNCAPLINSTLFTIEDILTSGAHSNPILMGVILGGTIVVVATAPLSSMALTA
 MLGLTGMPMAIGALS VFGSSFMNGVLFHKLKLGSRKDNIAFAVEPLTQADV
30 SANPIPIYVTNFVGGAACGILIALMKLVNDTPGTATPIAGFAVMFAYNPMIKVL
 ITALGCHLSLLAGYFGGIVF

Sequence description:

- 35 A] Length: 540 bp - 180 aa (partial sequence)
 B] N- and C-termini have yet to be elucidated

40

ID-143

Clone 2-c80

(SEQ ID NO: 131)
5 ATGTTTTTAAGTATAATGGCAGGTGTCATAGCATTTGTCCTGACAGTTATT
GCCATTCCACGCTTCATTAAGTTTTACCAATTGAAGAAAATTGGCGGGCAA
CAAATGCATGAAGATGTCAAACAACATCTAGCCAAAGCAGGTACGCCGAC
AATGGGAGGAACGGTATTTT
(SEQ ID NO: 132)
10 MFLSIMAGVIAFVLTVIAIPRFIKFYQLKKIGGQQMHEDVKQHLAKAGTPTMG
GTVF

Sequence description:

15 A] Length: 172 bp - 57 aa (partial sequence)
B] Shine Dalgarno sequence precedes 'ATG' start
codon. Possesses a potential leader peptide
sequence.

20

ID-144

25

Clone 3-83

(SEQ ID NO: 133)
30 ATGAAACCATATTTATCTTTTATTGGTAGAACGTTATTATACTTCGGTATTT
TATTGTTACTAATTTACTTTTTTGCATACCTTGGTCGCGGACAAGGCAGTTT
TATTTATAA
(SEQ ID NO: 134)
MKPYLSFIGRTLLEYFGILLLLIYFFAYLGRGQGSFIY

35

Sequence description:

40 A] Length: 113 bp - 37 aa (partial sequence)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. Possesses a
potential leader peptide sequence.
This orf is not in frame with nuc

ID-145

5

Clone 3-86

(SEQ ID NO: 135)
10 ATGTCATATTTTAGAAATTACTGGTATCGTTTTGGAGCAATTTTATTTATTA
TTTTAGCAGTAATATTGCTTGTTTTTAGACCTGACTGGTCAATGCTTCACTA
TCTATTGTATTTTTACTTTATGGCACTTCTAGCGCATCAATTTGAAGAATAT
CAGTTTCCCGGTGGGGGCATCACCTATCATTAACCTATGTTGTTTATGATGAA
GAAGAGCTGATGGATTGTTTTCCAGGCAATACTCAGTCTATTATGTTGGTT
AATACTATTGCTTGGTTGCTTTACATTGCTAGTATTGCTTTTCCTCAAGCTT
15 ATTGGCTTGGATTAGGAGTCATGTTCTTTAGTCTAACGCAGCTCTTGGGTC
ATGGTTTTTCAGATGAATATTAACTTAAACTTGGTATAATCCTGGTCTAG
CAACGACAGTATTTCTCCTAGTACCAATAGCTTGCGCATACATCTATCAAG
CTAGTGCAGAAGGAATGCTCACTTGGGGGAGATTGGCTAGGTGGTTTTATCA
TGTTGATTGTCTGTGTACTAACTAGCATTATTGCACCTGTACAGCTATTGAA
20 GGATAAGGAGACCAATTATATTATTAGTCCTTGGCAAATGGACCGTTTTCA
TAAGGTCGTTAATTTTGTAAGGATAAAAAAATAA

(SEQ ID NO: 136)
25 MSYFRNYWYRFGAILFIILAVILLVFRPDWSMLHYLLYFYFMALLAHQFEEYQ
FPGGASPIINYVVYDEEELMDCFPGNTQSIMLVNTIAWLLYIASIAFPQAYWLG
LGVMFFSLTQLLGHGFQMNILKKTWYNPGLATTVFLLVPIACAYIYQASAEG
MLTWGDWLGGFIMLIVCVLTSILAPVQLLKDKETNYIISPWQMDRFHKVVNFV
RIKK*

30 Sequence description:

35 A] Length: 651 bp - 219 aa (full length gene)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. Possesses a
potential leader peptide sequence.

40 ID-146

Clone 3-c88

(SEQ ID NO: 137)
ATGCCACTTACAGCACTTGAAATTAAAGATAAAACATTTTCATCAAAATTT
CGCGGTTATAGCGAAGAAGAAGTT

5 (SEQ ID NO: 138)
MPLTALEIKDKTFSSKFRGYSEEEV

Sequence description:

10

A] Length: 75 bp - 25 aa (partial sequence)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. No leader
peptide

15

ID-147

20 Clone 3-90

(SEQ ID NO: 139)
ATGTCACCTTTTCAAGAAAAAATTGCTTACAATTGCGCTAAAAAGGAAGCG
CTTTATAAAGAGAGTTTAGGACGCTACGCCTTGAGATCAATGCTAGCAGG
25 GGCTTATTTGACAATGAGTACTGCTGCCGGTATCGTCGCAGCTGATACTAT
TGGTAAAATTTCTCCTGCTCTATCAGGTTTTGTATTTGCTTTCATCTTTAGTT
TTGGACTTATTTATGTTTTAATATTTAATGGTGAATTGGCGACATCTAATAT
GCTTTATCTCACTGCAGGAGCCTATAATAAAAAATATCTCTTGGAAGAAAAGC
CATAACAATTTTAATTTATTGTACTTTTTTCAACCTCGTTGGTGCTTGTATA
30 TTAGCTTGGTTGTTTAA

(SEQ ID NO: 140)
MSLFQEKIAYNCAKKEALYKESLGRYALRSMLAGAYLTMSTAAGIVAADTIG
KISPALSGFVFAFIFSGLIYVLIFNGELATSNMLYLTAGAYNKNISWKKAITILI
YCTFFNLVGACILAWLF

35

Sequence description

40

A] Length: 406 bp - 125 aa (partial sequence)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. Possible
leader peptide

ID-148

5

Clone 3-92

(SEQ ID NO: 141)
10 AAGTTACAAGCGACTGAAGTTAAGAGCGTTCGGTAGCACAACCAGCTTC
AACAAACAAATGCAGTAGCTGCACATCCTGAAAATGCAGGGCTCCAACCTC
ATGTTGCAGCTTATAAAGAAAAAGTAGCGTCAACTTATGGAGTTAATGAA
TTCAGTACATAACCGTGCGGGAGATCCAGGTGATCATGGTAAAGGTTTAGC
AGTTGACTTTATTGTAGGTAAAAACCAAGCACTTGGTAATGAAGTTGCACA
15 GTACTCTACACAAAATATGGCAGCAAATAACATTTTCATATGTTATCTGGCA
ACAAAAGTTTTATTCAAATACAAATAGTATTTATGGACCTGCTAATACTTG
GAATGCAATGCCAGATCGTGGTGGCGTTACTGCCAACCCTATGACCACGT
TCACGTATCATTTAA

(SEQ ID NO: 142)
20 KLQATEVKSVPAQPASTTNAVAAHPENAGLQPHVAA YKEKVASTYGVNEF
STYRAGDPGDHGKGLAVDFIVGKNQALGNEVAQYSTQNMAANNISYVIWQQ
KFYSNTNSIYGPANTWNAMPDRGGVTANHYDHVHVSF

Sequence description

25

- A) Length: 419 bp - 139 aa (partial sequence)
- B) N- and C-termini have yet to be determined

30

ID-149

35

Clone 3-94

(SEQ ID NO: 143)
40 ATGATTCCAGTAGTTATTGAACAAACAAGTCGTGGTGAACGTTCTTATGAT
ATTTACTCACGTCTTTTAAAAGATCGTATTATTATGTTGACAGGCCAAGTT
GAGGATAATATGGCCAATAGTATCATTGCACAGTTATTGTTTCTCGATGCA
CAAGATAATACAAAGGATATTTACCTTTATGTCAATACACCAGGTGGTTCA
GTATCGGCTGGACTTGCTATTGTGGACACCATGAACTTCATTAAATCGGAC
GTACAGACGATTGTTATGGGGATGGCTGCTTCGATGGGAACCATTATTGCT
TCAAGTGGTGCTAAAGGAAAACGTTTTATGTTACCGAATGCAGAATATATG

5 ATCCACCAACCAATGGGCGGAACAGGCGGAGGTACACAGCAATCTGATAT
GGCTATCGCTGCTGAGCATCTTTTAAAAACGCGTCATACTTTAGAAAAAAT
CTTAGCTGATAATTCTGGTCAATCTATTGAAAAAGTCCATGATGATGCAGA
GCGTGATCGTTGGATGAGTGCTCAAGAACACTTGATTATGGCTTTATTGAT
GCTATTATGGAAAATAATAATTTACAATAATAGATTTAAAAGAGTTGAGTT
TACCAACTCTTTTTTTTATTTGTTGGAATTATGTTATAATCTTAGTAATTACA
GATATGACGCAGAAAGGAAAAAATTATTGA

(SEQ ID NO: 144)
10 MIPVVIEQTSRGERSYDIYSRLLKDRIIMLTGQVEDNMANSIIAQLLFLDAQDN
TKDIYLYVNTPGGSVSAGLAIVDTMNFIKSDVQTIVMGMAASMGTHASSGAK
GKRFMLPNAEYMIHQPMGGTGGGTQQSDMAIAAEHLLKTRHTLEKILADNSG
QSIEKVHDDAERDRWMSAQEHLIMALLMLLWKIIYNNRFRVEFTNSFFICW
NYVIILVITDMTQKGKNY*

15
Sequence description

20 A] Length: 693 bp - 231 aa (full length gene)
B] Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. No leader
peptide. Significantly, it would appear to have a
very hydrophobic C-terminus.

25
ID-150

30
Clone 2-c86

(SEQ ID NO: 145)
35 ATGAAACCAAAAaTTATTGGTGTACTTGGTCTAGGAATATTTGGACAAACA
CTCGCACAAGAACTAAGTAACTTTGAACAAGATGTTATTGCTATTGACAGC
AATCCTGAAAATGTACAAGCTGTCGCCGAAGT
TGTTACAAAAGCAGCTATCGGAGACATTACTGATTTAGCTTTCCTAAAACA
CATCGGGATCAGTGACTGTGATACTGTTATTATTGCTACAGGAAACAGTTT
AGAGAGCTCAGTATTGGCCGTAATGCACTGTAAAAAGTTAGGCGTCCCAC
AAGTTATTGCTAAAGCTCGAAACCTTGTATACGAAGAAGTACTTTATGAAA
TTGGTGCTGATTTGGTTATCTCTCCGGAGCGAGAATCTGGGGCAAAATGTTG
40 CTGCAAACCTCATGAGAAATAAAATTACAGATGTCTTCCAGATTGAATCTG
ATATTTCTGTCATTGAATTT

(SEQ ID NO: 146)
MKPKIIGVLGLGIFGQTLAQELSNFEQDVIAIDSNPENVQAVAEVVTKAAIGDI
TDLAFLKHIGISDCDTVIIATGNSLE

SSVLAVMHCKKLGVPQVIAKARNLVYEEVLVEIGADLVISPERESGQNVAAN
LMRNKITDVFQIESDISVIEF

5 Sequence description:

10 A] Length: 459 bp - 153 aa (partial sequence)
 B] Putative ATG start codon is preceded by a
 typical Shine-Dalgarno sequence. Possesses a
 potential leader peptide sequence.
 This orf is not in frame with nuc

15 ID-151

20 Clone 2-c88

 (SEQ ID NO: 147)
 GTGCGTTATAGTAAAGAGATTATTCAGTTAGCTATACCAGCTATGATTGAA
 AATATCTTACAAATGCTCATGGGAGTAGTTGATAATTATCTAGTGGCTCAG
 TTAGGTGTTGTAGCAGTATCAGGTGTTTCAGTTGCTAATAATAATTACT
25 ATTTATCAAGCTATTTTTATAGCTTTAGGGGCGAGTATAGCAAGTCTATTG
 GCCAAGTCGTTAGCAGGTAGTGAGAAGGATGATGCAATTTTCAGTATGTTCT
 CAAGCCATTTTTCTAACATCACTGATAGGGGCGAGTATTAGGAATTATCTCG
 ATTGTTTTTGGACAACTTTCTTT

30 (SEQ ID NO: 148)
 MRYSKEIIQLAIPAMIENILQMLMGVVDNYLVAQLGVVAVSGVSVANNIITIY
 QAIFIALGASIASLLAKSLAGSEKDDAISVCSQAIFLTSLIGAVLGIISIVFGQTFF

35 Sequence description

40 A] Length: 330 bp - 110 aa (partial sequence)
 B] Putative GTG start codon is preceded by a
 typical Shine-Dalgarno sequence. May have a
 leader peptide

ID-152

Clone 2-c92

5

(SEQ ID NO: 149)
TTGATTAAACAAGTATTCGTGCTTTTTGAAGAGGATTCTCCATAATAATACT
CCTTTAATAGTTATCGTGAGAAGTATTTTAAAGAAAAACCGCCAAGGTAG
AGCGACATTTCTGCCTTTAACTACAATAAAAACCAAGAGAATTAGCACAAAC
ATTATCTCTCAAAATTACAAAGTTCTCAAGGGTTTTTAGGAATAGCTAGTG
AATTGGTAACCTATGATCAACGCTTGTCAAACATTTTT

10

(SEQ ID NO: 150)
MINKYSCFLKRILHNNTPLIVIVRSILKKNRQGRATFLPLTTIKPRELAQHLYSK
LQSSQGFLGIASELVTYDQRLSNIF

15

Sequence description

20

A] Length: 240 bp - 80 aa (partial sequence)
B] No obvious Shine Dalgarno sequence precedes the Putative TTG start
codon

25

ID-153

Clone 2-c94

30

(SEQ ID NO: 151)
TTGTTGACTCACAAAAATATATTATTAACCATTATATTTGGATTATTTATGA
TTATATTATCAGCATGTGGTATGTCTAATAAGGAAATGGCTGGTATTGATA
ATTGGGAACATTATCAAAAGGAAAAGAAAATTACTATTGGATTTGATAAT
ACTTTTGTTTCCTATGGGATTTGAAAGTCGTTCTGGTGACTATACCGGCTTTG
ATATTGATTTAGCTAATGCTGTTTTTAAAGAATACGGTATTTTCAGTGAAAT
GGCAGCCTATTAACCTGGGATATGAAAGAACTGAACTTAATAATGGTAAT
ATAGACCTTATTTGGAATGGTTATTCAAAAACGGCAGAACGTGCTAAAAA
AGTCGCTTTTACAAACCCATATATGAATAATCATCAAGTAATTGTTACTAA
AACTTCATCACATATTAATAGTATTAAGGATATGAAGGGGAAAAAACTAG
GAGCCCAGTCGGGTTTCATCTGGTTTTGATGCTTTTAACGCTAAACCTGATA
TTTTAAAAAAGTTTGTAAAAGGAAAAGAAGCAGTTCAATACGATACTTTC
ACTCAGGCTTTGATTGATTTAAAAAATAACCGTATTGATGGTCTTTTGATT
GATGAAGTTTATGCTAACTATTATTTAAAGCAAGAAGGAA

40

(SEQ ID NO: 152)
MLTHKNILLTIIFGLFMILSACGMSNKEMAGIDNWEHYQKEKKITIGFDNTFV
PMGFESRSGDYTGFDIDLANAVFKEYGISVKWQPINWDMKETELNNGNIDLI
WNGYSKTAERAKKVAFTNPYMNNHQVIVTKTSSHINSIKDMKGKKLGAQSG
SSGFDAFNAKPDILKKFVKGKEAVQYDTFTQALIDLKNNRIDGLLIDEVYANY
YLKQEG

Sequence description

10 A] Length: 649 bp - 216 aa (partial sequence)
B] TTG start codon is preceded by a possible
typical Shine-Dalgarno sequence. Has a
leader peptide

ID-154

20 Clone 2-c100

(SEQ ID NO: 153)
25 ATGAAAATTTGGAAAAAATAACCTTAATGTTTTCTGCAATTATTTTAACA
ACAGTAATTGCATTGGGAGTCTATGTTGCCTCAGCTTATAATTTTTCGACTA
ATGAATTGTCTAAGACTTTT

(SEQ ID NO: 154)
MKIWKKITLMFSAILTTVIALGVYVASAYNFSTNELSKTF

30 Sequence description

35 A] Length: 123 bp - 41 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Has a
typical leader peptide

40 ID-155

Clone 2-c1

(SEQ ID NO: 155)
ATGAAAAACAAAGACTATTACTGCTTTTTGGAGGCTTATTAATAATGATA
ATGATGACAGCATGTAAGGATTCAAAAATCCCAGAAAACCGCACGAAAAA
5 GGAATACCAGGCAGAACAGAATTTTAAGTCATACTTTAAATATATATCAG
ATAAAAATAACTATTTAGATAATATAAAAGTTTATTACTTTTCTATAAGTA
TTTCTAAAGATGTACAAGATAAAGTCAGTGAAACAACAACCTTGTTTCATATA
GACTAGAAAAGCAAAAAGAATCAAGAGTTCATTGGTAATTTTGAACATGAA
GTTAGTGAATCTAGTCAATATTCAACCGAAGTTAAAAATCAAATACAGTAT
10 CCAATCCAGTATAAAGATAATTCAATTCGTTTTACTGAAAAAACACCGTCA
GAACGTTATGATGAGTTTGTGTTTTAGTTTCATTTGATTCTTCATTATTA
AATATAAAATATATGATTACTTACTAAAACATCCCGAAACTGAATTAAAA
GGTGTTTCCTATAAGATTCCTATAAATTCTGAAATTGTAGCCCCTTTTATAA
ATCAATTAAATATAAAAAATCCTAAAAAATCATCTATTTCGGTTACAAAAA
15 CGGAAAGTAAAGAATATTATTATACAATCAGTATTGATACTGATTCTGAGA
TATATTCTATATTCGAAGGTATTCAT

(SEQ ID NO: 156)
MKKQRLLLLFGGLLIMIMMTACKDSKIPENRTKKEYQAEQNFKSYFKYISDKN
NYLDNIKVYYFSISISKDVQDKVSETTTCYSRLEKQKNQEFIGNFEHEVSESSQ
20 YSTEVKNQIQYPIQYKDNSIRFTEKTPSERYDEFVFSSFDSSLLKKYKIYDYLLK
HPETELKGVSYKIPINSEIVAPFINQLNIKNPKKSSISVTKTESKEYYYTISIDTDS
EIYSIFEGIH

25 Sequence description

30 A] Length: 687 bp - 229 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Has a
typical leader peptide. C-terminus has yet to be
verified

35 ID-156

Clone 2-c5

40 (SEQ ID NO: 157)
ATGACATTTGACACCATTGATCAATTAGCGGTTAATACAGTCCGCACGCTT
TCTATTGATGCTATCCAAGCAGCAAATTCTGGGCACCCAGGTCTTCCTATG
GGAGCTGCGCCTATGGCTTATGTGCTTTGGAATAAATTCTTAAATGTAAAC
CCAAAAACAAGTCGCAATTGGACAAACCGTGACCGTTTTGTACTTTCAGCT

GGGCATGGTTCAGCTCTTCTTTATAGCCTACTTCATTTAGCTGGCTATGATT
TATCAATTGATGATTT

(SEQ ID NO: 158)
MTFDTIDQLAVNTVRTLSIDAIQAANS GHPGLPMGAAPMAYVLWNKFLNVNP
5 KTSRNWTNRDRFVLSAGHGSALLYSLHLAGYDLSIDD

Sequence description

10

A] Length: 272 bp - 90 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. No obvious
leader peptide

15

ID-157

20

Clone 2-c8

(SEQ ID NO: 151)
ATGAGAACTACTATTTAGAAATGATATTTGCTATTCCAAAGTTTATCTTTAGA
25 TTGATTTGGAATATCATTGTTGGGGAATATTCAAGACAGTTCTTGTTATTGCG
ATTATTTTATTTGGCTTGTATTACTATGCGAATCACAGTCAATCAGAATTG
CTAATCAACTTAGTGACATTATTCAGACAGGAAAAACATTTT

(SEQ ID NO: 160)
MRTLFRMIFAIPKFIFRLIWNIIWGIFKTVLVIAIILFGLYYYANHSQSEFANQLS
30 DIIQTGKTF

Sequence description

35

A] Length: 197 bp - 65 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Possesses a
leader peptide

40

ID-158

Clone 2-c9

5

(SEQ ID NO: 161)
ATGTCAAAAAAAAAATAATATTAGGAATTTTATCTCTTTTATCTGTCGTTACTT
TGGTGGCGTGTGGTTCATCAGACAAACAGCTACAAGATAAAGTTGAGAAA
AAAGGGAAGTTAGTTTTAGCGGTGAGTCCAGATTATGCTCCCTTTGAGTTT

10

(SEQ ID NO: 162)
MSKKIILGILSLLSVVTLVACGSSDKQLQDKVEKKGKLVLA VSPDYAPFEF

Sequence description

15

A] Length: 153 bp - 51 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Possesses a
leader peptide (not in frame with nuc)

20

ID-159

25

Clone 2-c10

30

(SEQ ID NO: 163)
ATGAAAAATCAAAGACTATTACTGCTTTTTTGGAGGCTTATTAATAATGATA
ATGATGACAGCATGTAAGGATTCAAAAATCCCAGAAAACCGCACGAAAAA
GGAATACCAGGCAGAACAGAATTTTAAGTCATACTTT

35

Sequence description

40

A] Length: 139 bp - 46 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Possesses a
leader peptide

(SEQ ID NO: 164)
MKNQRLLLLFGGLLIMIMMTACKDSKIPENRTKKEYQAEQNFKSYF

ID-160

Clone 2-c11

5

(SEQ ID NO: 165)

ATGATTGGAAAATTATATTATAGCTATAGAAAGTCACGCTTATTAAGAAGT
ATTTTATGGCTTATTTTAATTGTTGGTGTATATATGTTAGGACAACGTGTTT
TATTATCCACTGTTTCCTTTATCACATCAAGAGATAAAACTAGCAGTAGATC
10 AACATTTACTCAATAACTTTTCAGCAGTAAGTGGTGGGAGTTTTAATAAAT
TAAATGTTTTTCACACTGGGGTTGAGTCCATGGATGTCAAGTATGATTATTT
GGAGATTCGTTTCCTTATTTTCGTGGGCAAAAAATGCAACGAAGCGAAAA
GCAGAAGTAGCTCAATATACTTTAATGCTTACTATCTCAGTTATACAAGCA
TATGGTGTTCAGGAAATCAATTTATAAAAAGCTCTTTATTAGGTTCTTATA
15 GTGATATTGTTTTT

(SEQ ID NO: 166)

MIGKLYYSYRKSRLRSILWLILIVGVYMLGQRVLLSTVPLSHQEIKLAVDQHL
LNNFSAVSGGSFNKLVFTLGLSPWMSSMIWRFVSLFSWAKNATKRKA EVA
QYTLMLTISVIQAYGVSGNQFIKSSLLGSYSDIVF
20

Sequence description

25 A] Length: 423 bp - 141 aa (partial sequence)
B] ATG start codon is preceded by a potential
typical Shine-Dalgarno sequence. Possesses a
leader peptide

30

ID-161

35 Clone 2-c13

(SEQ ID NO: 167)

ATGAAAGGTCTATTGGATTTTTTTAGTTAATATTGCCAGAACGCCAGCTATT
TTAGTCGCCTTGATAGCCATTATCGGTTTAGTACTGCAGAAAAAAGGTGTT
40 CCTGATATTGTAAAAGGTGGAATAAAAAACATTTGTTGGCTTCTTAGTGGTT
TCTGAAGGTGCAGGGATAGTCCAAAATTCCTTGAATCCATTTGGAAAAATG
TTTGAACATGCTTTTTCATTTGGTGGGGGTAGTTCCTAATAATGAAGCCATT
GTAGCAGTAGCTCTTACGAAGTATGGCTCAGCAACTGCTTTGATTATGTTA
GCGGGAATGATTTTTTAATATTTTAATTGCTCGTTTTACAAAA

(SEQ ID NO: 168)
MKGLLDFLVNIARTPAILVALIAIIGLVLQKKGVDPDIVKGGIKTFVGFLVVSEG
AGIVQNSLNPFGKMFHAFHLVGVVPNNEAIVAVALTKEYGSATALIMLAGMI
FNILIARFTK

5

Sequence description

- 10 A] Length: 348 bp - 116 aa (partial sequence)
 B] ATG start codon is preceded by a potential
 Shine-Dalgarno sequence. Possible leader
 peptide

15

ID-162

Clone 2-c21

20

(SEQ ID NO: 169)
TTGGTTGGTAAGCCCCAATTACTATTTTTAGATGAACCTACTTCCGGAATG
GATACTTCCACACGTCAACGATTTTGGGAAGCTGGTTGCGACACTAAAAAA
AGAAGGTGACACAATTGTCTATTCTAGTCATTATATCGAAGAGGTAGAAC
25 ATACAGCTGATAGGATTTTAGTACTTCATAAAGGAAAGTTATTACGCGATA
 CAACCCCCTTTGCCATGAAGCAAGAAAAAAACCGAAAAGTTATTCACCGTT
 CCGCTTAGTTATCAAAAATTATTACCTACCTATTTGATTACAGAGTGTGAA
 GCCAAGAGTGATAGTATAACGTTTGTACTGGGGAGGCTGAAACTGTATG
 GAAAATACTGGCAGATAATGGTTGTCCTATTGAAGCTATTGAGATGACCA
30 ATAGAACTTTGTAAATCGTATTTTTGAGACTACTAAGGAGGTAAACATG
 AGAATCTTTA

30

(SEQ ID NO: 170)
MVGKPQLLFLDEPTSGMDTSTRQRFWKL VATLKKEGDTIVYSSHYIEEVEHTA
DRILVLHKGKLLRDTTPFAMKQEKEKLFTVPLSYQKLLPTYLITECEAKSDSI
35 TFVTGEAETVWKILADNGCPIEAIEMTNRTLNRIFETTKEVKHENL

35

Sequence description

40

- A] Length: 462 bp - 155 aa (partial sequence)
B] B] Putative TTG start codon is not preceded by
an obvious Shine-Dalgarno sequence. No obvious
leader peptide. N- and C- termini require further



examination.

ID-163

5

Clone 2-c25

(SEQ ID NO: 171)
10 TTGAAAAAATCCCAAGAGAAGCCGTAAGGCAGTGACAACAAGTGGTGAGA
AGACTTTACTTGAGGATTTGGCAAAAATGAATTTCTAGACGAAGTCATTA
ATGTTATGGTTTTATATACCTTGAATAAGACAAAATCTGCTAACTTAAATA
AGGCCTATATCATGAAAGTTGCTAATGATTTTGCCTTTCAGAATGTTATGA
CGGCCGAAGATGCTGTGCTTAAAATTCGTGATTTTTCAGATCAAAAAGTAA
GGACTAAAACAGAAACGAAGAAGAAACAATCGAATGTTCTGAATGGAGT
15 AATCCTGATTATAAAGATGAGGTTAGCCCAGAAAAAGAAATTGAATTAGA
ACAGTTT

(SEQ ID NO: 172)
20 MKKSKRSRKAVTTSGEKTLLEDLAKMNFLEVINVMVLYTLNKTKSANLNK
AYIMKVANDFAFQNVMTAEDAVLKIRDFSDQKVRTKTETKKKQSNVPEWSN
PDYKDEVSPKEIELEQF

Sequence description

25

A] Length:360 bp - 120 aa (partial sequence)
B] N- and C- termini require verification.

30

ID-164

Clone 2-c28

35

(SEQ ID NO: 173)
ATGACGAATCATATTACTAACTGATAGAAAATAGCGGAAAAAAATTGAC
AGAAATTAGCGAAGCTACAGATATAGCCTATCCTACACTTTCTGGATACAA
TCAAGGAATCCGCAAACCTAAAAAAGATAATGCTGAAAAATTGGCAAAAT
40 ACTTTAATGTTTCCGTCGCTTACATTATGGGACTTGATAGCAACCCACATG
CTCCATCAAATCTT

(SEQ ID NO: 174)
MTNHITKLIENSGKKLTEISEATDIAYPTLSGYNQGIRKPKKDNAEKLAKYFNV
SVAYIMGLDSNP HAPSNL

Sequence description

5

A] Length:218 bp - 72 aa (partial sequence)
B] ATG start codon is preceded by an
obvious Shine Dalgarno sequence. No obvious
leader peptide.

10

ID-165

15

Clone 2-c29

(SEQ ID NO: 175)
20 TTGATGAAAAGGAATAAACATTTACCGTTAACAGAACTACCTATTATATT
TTATTAGCTTTGTTTGAGGAAGCGCATGGCTATTATGAAAAAAGTT
GAAGAAATGAGTGGCGGTGATGTTAGAATAGCCGCAGGGACAATGTACGG
TGCCATTGAAAATTTACTTAAACAAAAAATGGATAAAGTCTATCTCAAGTGA
CGATAGAAGAAGAAAAGTTTATATTATTACTGAGACAGGAAAAGAAATAG
TAGAACTTGAAACGAATCGATTAAGAAAGTTACTTAATACTGCTAATCAGT
25 TGGGTTTTGGAGGAGATGGTTATGATAAAGTTT
(SEQ ID NO: 176)
MMKR NKHLPLTETTYIILLALFEEAHGYAIMKKVEEMSGGDVRIAAGTMYG
AIENLLKQKWIKSISSDDRRRKVYIITETGKEIVELETNRLRKLLNTANQLGFG
GDGYDKV

30

Sequence description

35

A] Length:337 bp - 112 aa (partial sequence)
B] TTG start codon is preceded by an
obvious Shine Dalgarno sequence. Actual start
codon may ATG that comes immediately after the
TTG. Potential leader peptide.

40

ID-166

Clone 2-c35

(SEQ ID NO: 177)
5 CCCATTACTGGTGAGTTAATAGCTGAGAAATTAGGAGTACCAAGAGCAGC
ACTAAGGTCTGATTTGCGGGTTTTAAGTATGCTAGGTATCATAGATGCAAA
ACCTAAGGTTGGTTATTTTTATTTAGGACAGTATCATGCTTCAATAGGGAC
AAGTCATTTTGAAAAGATGACAGTTTCAGAAATTATGGGGATCCTTCTGAC
AGTTCATCAAAAAGATTCAGTTTATGATGTTATTGTACATATTTTTATGGA
10 AGATGCTGGTTGTGCTTTTATCTTGGATGATGATGATTTTCTCTGTGGAGTC
GTGTCACGTAAAGATTTACTAAAAACCAGTATTGGCGGAGGAGATCTTTCT
AAAATGCCAATAGGAATGGTGATGACACGTATGCCACACGTGACAACGT
TTTAGAAAATGAAAGTCTTTTTGCGGCAGCTGATAAATTAGTGAGCAGAA
AAGTGGATAGTCTCCCTGTCGTTTCGTCATGATAAGCAATATCCCGAAAAAT
15 TTA
(SEQ ID NO: 178)
PITGELIAEKLGVPRALRSDLRVLSMLGIIDAKPKVG YFYLGQYHASIGTSHF
EKMTVSEIMGILLTVHQKDSVYDVIVHIFMEDAGCAFILDDDDFLCGVVSRKD
LLKTSIGGGDLSKMPIGMVMTRMPHVTTVLENESLFAAADKLVS RKVDSL PV
20 VRHDKQYPEKF

Sequence description

25 A] Length:511 bp - 170 aa (partial sequence)
B] N- and C-termini to be determined

30 ID-167

Clone 2-44

35 (SEQ ID NO: 179)
TTGGAAGTCATCATGCAATTTATTTATAGTATTATTGGTATTTTATTGGTAT
TAGGAATTGTGTATGCAATTTCTTTCAATCGTAAGAGTGTTTCTCTAAGTTT
AATTGGAAAAGCTCTTATCGTTCAATTCATTATTGCGCTAATCTTAGTACGT
40 ATCCCACTAGGCCAACAAGTTGTTAGTGTTGTTTCAACTGGAGTTACTAAA
GTAATCAACTGTGGTCAAGCTGGTTT
(SEQ ID NO: 180)
MEVIMQFIYSIIGILLVLGIVYAI SFNRKSVSLSLIGKALIVQFI ALILVRIPLGQQ
VVS VVSTGVTKVINCGQAG

Sequence description

5

A] Length:233 bp - 77 aa (partial sequence)
B] TTG start codon is preceded by a
possible Shine Dalgarno sequence. Actual start
codon may occur further downstream. Potential
leader peptide.

10

ID-168

Clone 2-46

15

(SEQ ID NO: 181)
CAACCTAATAAAGCTTTAGAAAGTGATGAGATTGATATTAATGCTTTCCAG
CATTATAATTACTTAACCAATTGGAATAAAGCAAATAAGACCAATCTTGTT
TCCGTTGCTGAGACATACTTTACTTCCTTTAGATTATACTCTGGTACTAAGA
ACGGTAAAGGTAAATACCAAACAGTTTCTGAAATTCCAAATAAAGCAACT
ATTACTATCCCAAACGATGCAGTTAACGAAAGTCGCTCTCTCTACTTGTTA
CAATCAGCAGGCTTGCTAAAATTGAAAGTATCAGGTGATACATTAGCAAC
AATGTCAGATGTTGTTTCCAATCCTAAATCTTTAGATTT

20

25

(SEQ ID NO: 182)
QPNKALESDEIDINAFQHYNLTNWNKANKTNLVSVAETYFTSFRLYSGTKN
GKGKYQTVSEIPNKATITIPNDVNESRSLYLLQSAGLLKLKVSGDTLATMSD
VVSNPKSLD

30

Sequence description

A] Length:344 bp - 114 aa (partial sequence)
B] N- and C- termini require verification

35

ID-169

40

Clone 2-47

(SEQ ID NO: 183)
ATGAAATGTATAATAAATAATATAAATAAAATAAAAATGATAATTGAGAT
TTATCATAGAAGGAAAACATTTTGAATTAATAAAAATCATATTATCTAC

TGCAGCTCTTACTGCTCTCTTTTTAGGATATAATAGCGTTACTGCGGATACA
TATAATAACTATCAGCCACATAGATCAAATAATATGGATTTAAGTGAAGGA
ATATAACTATAATAACCAGATAGAACTTCAGGAGCGTATAAAAAACCTAA
ATATACCTTTT

5

(SEQ ID No: 184)
MKCIINNINKIKMIEIYHRRKTILKLNKIILSTAALTALFLGYNSVTADTYNNY
QPHRSNNMDLTEEYNYNQIELQERIKNLNIPF

10

Sequence description

A) Length: 264 bp - 88 aa (partial sequence)

B) There is a Shine-Dalgarno sequence upstream
of this sequence. Potential leader peptide
sequence

15

20

ID-169

Clone 2-47

25

ATGAAATGTA
TTATCATAGAAG
TGCAGCTCTTACTGC
TATAATAACTATCAGCC
ATATAACTATAATAACCAGATAGAACTTCAGGAGCGTATAAAAAACCTAA
ATATACCTTTT

30

MKCIINNINKIKMIEIYHRRKTILKLNKIILSTAALTALFLGYNSVTADTYNNY
QPHRSNNMDLTEEYNYNQIELQERIKNLNIPF

35

Sequence description

40

A) Length: 264 bp - 88 aa (partial sequence)

B) There is a Shine-Dalgarno sequence upstream
of this sequence. Potential leader peptide
sequence

ID-170

5 Clone RS-58b

(SEQ ID NO: 185)
TTGGGTGATTATTATGGTAAGAAATATTTTGGTGAGGCAGCTAAAAAAGA
CGTCGAACATATGGCTAAGAAAATCATTAAATGTCTATAAAACACGGTTAA
10 AAAACAACACTTGGTTATC
AGAAAATACAAAAGCAATGGCCATTAAGAAACTTGATAACATGAGATTAA
TGATTGGCTATCCAGAAGATTATCCTGATCTTTATCGTCAGTACCAATTTG
ATAGTAAAGCAAGCTTCTTTGAAAACAATGATAACTACAGAAAATTATCG
AACAGAAAACATTTGAAGAATTAAACCAGTCTAATCAACGTGAACATTG
15 GCAAATGAGTGCCAATGCTGTAAATGCTTATAATGATCCTAATACCAATTC
CATAGTCTTTCCAGCAGCGATTTTCAATCACCAGTGTACGATAAACTAA
AACAGTTAGTCAAAATTATGGAGCTATCGGAGCAATTATTGGTCATGAAAT
TTCACACTCATTTGATATTAATGGTATGAAATATGACGAGAAAGGGAATCT
TCACGATTGGTGGACTAAAGAAGATTTAAATCATTATAAGAAATCAACAC
20 AAGCTATGATTGACCAATGGGATGGCCTTAAAGCAGATGGCGGTAAAGTT
GATGGTAAATTAACCTTTAGCAGAAAATATTGCAGATAATGGTGGTGTATG
GCATCTCTAGAAGCTCTTAAGACTGAAAAAATCCAACTATAAAGAATTTT
TTGAATCATGGGCAAGTATTTGGCGTCAAAAAGCAACCAAAGAACAAAGT
AAGTCCTCAATTCAGTCAGATGTTTCATGCACCATATGAATTGA
25 GAGCTAACATCCCAGTACGTAATTTCCAAGAATTTTATGATGCCTTTGGTG
TTAAAAAAGGCGATTCAATGTATCTAAAACCAGAAAAACGTTTGACACTTT
GGTAA

(SEQ ID NO: 186)
30 MGDYYGKKYFGGEAAKKDVEHMAKKIINVYKTRLKNNTWLSENTKAMAIKK
LDNMRLMIGYPDYPDLRQYQFDSKASFFENNDNYRKLSNKKTFEEFNQSNQ
REHWQMSANAVNAYNDPNTNSIVFPAAIFQSPLYDKTKTVSQNYGAIGAIIHG
EISHSFDINGMKYDEKGNLHDWWTKEDLNHYKKSTQAMIDQWDGLKADGG
KVDGKLTLAENIADNGGVMASLEALKTEKIQTIKNFLNHGQVFGVKKQPKNK
35 VSPQFSQMFMHHMN*

Sequence description:

40 A] Length: 819 bp - 272 aa (full length gene)
(107 bp of additional DNA sequence (> onwards) is
also included. While not in-frame with the
described orf, it also shares strong homology
with the neutral peptidases.

→
B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-89 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-89 gene sequence. ID-89 and ID-170 together show homology over their combined entire length with the neutral endopeptidases from *Lactococcus* and *Lactobacillus*. Possesses TTG (possible ATG start codon located 13 bp further downstream) start codon with no obvious signal peptide. Shine Dalgarno sequence not immediately obvious. Possibly located further downstream

10 ID-171

Clone 2-18/22b (Mod2)

15 (SEQ ID NO: 187)
ATGACCATGATTACGCCAAGCTTCATTAAGGTATCTCTAGATGAAACAAAT
CGTATGATGCGTATGATATCAGATTTATTAAGTTTATCGCGCATTGATAAT
GAAGTAACGCATTTAGATGTTGAAATGACGAATTTTACAGCTTTCATGACC
TCAATTTTGAATCGATTTGATCAGATTAGAAATCAAAAAACAGTCACAGG
20 AAAAGTTTATGAAATTGTCAGAGATTATCCTCTTAAGTCAATTTGGGTGGA
AATTGATACAGATAAGATGACTCAAGTGATTGATAACATTTTAAATAATGC
AGTCAAGTATTCACCAGATGGTGGTAAGATTACAGTTAATCTACGCACAAC
TAAAACGCAGATGATTTTATCAATATCAGACCAAGGCTTAGGTATTCCCAA
AAAAGATTTACCTCTCATTTTTGATCGTTTTTATCGTGTTGATAAGGCGAGA
25 AGTCGTCAACAGGGTGGGACTGGACTTGGTTTGTCAATTGCAAAAGAAAT
TGTTAAGCAGCATAAGGGATTTATTTGGGCTAAGAGTGAGTATGGTAAAG
GGTCTACTTTTACAATCGTCTTGCCTTATGATAAAGATGCTGTAACCTTATGA
AGAATGGGAGGACGTTGAAGATTAA

30 (SEQ ID NO: 188)
MTMITPSFIKVSLDETNRMMRMISDLLSLSRIDNEVTHLDVEMTNFTAFMTSIL
NRFDQIRNQKTVTGKVYEIVRDYPLKSIWVEIDTDKMTQVIDNILNNAVKYSP
DGGKITVNLRTTKTQMILSISDQGLGIPKKDLPLIFDRFYRVDKARSRQQGGTG
LGLSIAKEIVKQHKGFIWAKSEYGKGSTFTIVLPYDKDAVTYEEWEDVED*

35

Sequence description:

40 A] Length: 613 bp - 212 aa (full-length gene possibly)
B] Possible Shine Dalgarno sequence present
upstream of a ATG start codon. May not have yet
determined the N- portion of this gene. No
obvious signal peptide.

ID-172

5

Clone 2-54balternate (107b)

(SEQ ID NO: 189)
10 TTGAAAAAAATTATTACTTCTATTCTATTACTTAGTTGCATTTTTTTTATGC
CAACCATCTCTGCTGAATCTTTTAATGCTTCCGCTAAACATGCCTTAGCAGT
TGATTTAGATTTCAGGAAAAATCTTGTATGAAAAAGATGCTAACAAACCCG
CTGCTATTGCTTCCTTGACTAAAATAATGACCGTTTATATGGTCTATAAAG
AAATTGATAACGGTAACCTCAAGTGGAATACCAAAGTAAATATATCTGAC
15 TACCCTTATCAACTAACACGCGAATCTGATGCTAGTAATGTTTCCTTTAGAA
AAAAGGCGCTATACTGTAAACAACCTCGTGGACGCTGCCATGATTTCTAGT
GCTAACAGTGCAGCCATTGCTTTAGCTGAACATATTTTCAGGAACTGAAAGT
AAATTTGTTGATAAAAATGACTGCTCAATTGGAAAAGTGGGGAATTCATGAT
AGCCACCTAGTCAATGCTTCTGGCTTAAATAATAGTATGTTAGGCAATCAC
20 ATTTATCCAAAATCGTCACAAAACGACGAAAATAAAAATGAGTGCACGTGA
TATTGCTATTGCTGCCTACCATTTGGTCAACGAATATCCTTCCATTCTTAAG
ATTACTAGTAAGTCCGTTGCTAAATTTGATAAAGATATTATGCATTCTTAT
AACTACATGCTACCAGATATGCCTGTCTTTAGACCAGGTATTACAGGTTTG
AAAACCTGGGACAACGGAATTAGCTGGCCAATCTTTTATTGCTACATCTACT
GAAAGTGGAATGAGACTACTCACTGTTATTATGCATGCTGATAAGGCCGAT
25 AAAGACAAATATGCTCGCTTTACAGCAACTAACTCTCTCTTGAACATATC
ACAAACACCTACGAACCTAACCTTGTATTAGCTAAAGGAGCTGCATATAA
AGGTAAAGAAGCAAGTGTGAGAGACGGAAAAGAACAATCGGTCATCGCT
GTTGCTAAAAACGATTTGAAAGTAGTACAGAAGAAAAATATCACTAAACA
AAATCAGTTAAAAAATTAACCTTTAAAAAAGAGCTTACTGCTCCTATTACAAA
30 AAAAGAGAACCTAGGGAAAGCTTATTACGTTGACCTTAATAAGGTTGGAA
AAGGCTATCTCATAAAGGAACCTAGCGTTCATTTAGTGGCAAAGATAGT
ATTGAGCGCAGTTTCTTCCTCAAAGTGTGGTGGAATCATTTTGTGCGCTAC
GTTAACGAAAAACTTTAA

35

(SEQ ID NO: 190)
MKKIITSILLSCIFFMPTISAESFNASAKHALAVDLDSGKILYEKDANKPAAIA
SLTKIMTVYMVYKEIDNGNLKWNTKVNISDYPYQLTRESNASNVPLEKRRYT
VKQLVDAAMISSANSAAIALAEHISGTESKFVDKMTAQLEKWGIHDSHLVNA
SGLNNSMLGNHIYPKSSQNDENKMSARDIAIAAYHLVNEYPSILKITSKSVAKF
40 DKDIMHSYNYMLPDMPVFRPGITGLKTGTTELAGQSFIATSTESGMRLLTVM
HADKADKDKYARFTATNSLLNYITNTYEPNLVLAKGAAYKGKEASVRDGKE
QSVIAVAKNDLKVVQKKNITKQNQLKINFKKELTAPITKKENLGKAYYVDLN
KVGKGYLIKEPSVHLVAKDSIERSFFLKVWWNHFVRYVNEKL*

Sequence description:

A) Length: 1236 bp - 412 aa (full-length gene sequence possibly)

B) A possible Shine-Dalgarno sequence precedes the putative 'TTG' start codon. (needs further cloning and sequencing to verify N-terminus)

ID-173

Clone 3-60b

(SEQ ID NO: 191)

ATGACGCTTCGAGAATTAACAATAGAAGAATTTAAAGAACATTCAGGAAA
TTATGATTCACAATCATTTTTACAAACACCTGAGATGGCTAAACTTTTAGA
AAAACGCGGCTATGATGTTAGGTATTTGGGATATCAAGTAGAAAATAAAC
TAGAGATAATCAGTTTATCTTATATTATGCCAGTCACTGGTGGTTTTCAAAT
GAAAATTGATTCAGGACCAGTTCATTCAAATTCTAAGTATCTAAAACAATT
TTATAAAGCATTGCAAGGCTATGCCAAATCCAACGGTGTTCTAGAATTAAT
AGTTGAGCCTTTTGATGATTACCAATTATTCAGTTCGGGAGTTCCTAGT
AATCAGGGGAAATGATAATCTGATTGAAGATTTTACCAGTTCAGGTTATCAC
CATGATGGTTTAAACAACCTGGTTTTACTGGTAAATATTTATCTTGGCACTATG
TTAAAAATTTAGAAGGTGTCACCTTCTGAAACGTTACTATCTTCATTCTCTAA
GACAGGACGAGCTTTGGTTAAGAAAGCAATGTCTTTTGGAATCAAGGTTC
GCGTTCTTAAACGTGATGAGCTACATTTATTTAAAGAGATAACAACCTTCTA
CGTCAAATAGACGTGATTATATGGATAAGTCCTTAGATTATTATCAAGATT
TTACGATAGCTTTGAAGGCAAGGCTGAATTTGTGATTGCCACTTTAAATT
TTAGAGAATACGACCATAACTTGCAAATAAAAGCTGAAGCATTGGAAAAT
AAGCTT

(SEQ ID NO: 192)

MTLRELTIEEFKEHSGNYDSQSFLQTPEMAKLLEKRGYDVRYLG YQVENKLEI
ISLSYIMPVTGGFQMKIDSGPVHSNSKYLKQFYKALQGYAKSNGVLELIVEPF
DDYQLFTSSGVPSNQGNLIEDFTSSGYHHDGLTTGFTGKYL SWHYVKNLE
GVTSETLLSSFSKTGRALVKKAMSFGIKVRVLKRDELHLFKEITTSTSNRRDY
MDKSLDYYQDFYDSFEGKAEFVIATLNFREYDHN LQIKAEALENKL

Sequence description

A) Length: 771 bp - 257 aa (partial gene sequence)

B) This gene sequence was not identified using the LEEP system. It was identified immediately downstream of the ID-65 gene which was identified by

LEEP, during cloning and sequence analysis of the full-length ID-65 gene sequence. Sequence Characteristics:
No obvious leader peptide sequence
Orf is preceded by a potential Shine-Dalgarno sequence.

ID-174

Clone 2-17b (ID-80b)

(SEQ ID NO: 193)
TTGTCATTAAGTTTGGTTGCAGTGTTAAATCTTATCCCTCCTAAAATCATGG
GATCAGTTATTGATGCTATTACAACCTGGAAAATTAACAAGACCACAATTAC
TATGGAATTTATTAGGTTTGGTTTTTGTTCAGCTTTAGCTATGTATGGGCTGCG
TTATATTTGGCGTATGTATATTTTAGGGACTTCTTACAAATTAGGCCAAGTT
GTCAGATACCGTTTATTTGAACATTTTACAAAAATGTCTCCTTCTTTTTATC
AGAAATATCGTACAGGTGATTTAATGGCGCACGCGACCAACGACATCAAT
TCTCTAACACGTCTTGCAGGAGGAGGAGTTATGTCAGCAGTGGATGCCTCT
ATCACAGCATTAGTAACGCTTATCACCATGTTCTTTACTATTTCGTGGCAA
ATGACATTAATTGCGGTTATCCCTTTGCCCTTAATGGCCTTAGCACTAGTA
AATTGGGGCGAAAAACCCATGAAACCTTCAAAGAATCTCAGGCAGCCCTT
TTCAGAATTAAATAATAAAGTG

(SEQ ID NO: 194)
MSLSLVAVLNLIPPKIMGSVIDAITTGKLTRPQLLWNLLGLVLSALAMYGLRYI
WRMYILGTSYKLGQVVRYRLFHFTKMSPSFYQKYRTGDLMAHATNDINSLT
RLAGGGVMSAVDASITALVTLITMFFTISWQMTLIAVIPLPLMALALVNWGEK
PMKPSKNLRQPFSELNNKV

Sequence description

A) Length: 534 bp - 178 aa (partial gene sequence)

B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-80 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-80 gene sequence.

Sequence Characteristics:

No obvious leader peptide sequence
Orf is preceded by a potential Shine-Dalgarno sequence.

ID-175

5 Clone 2-11Ab (ID-103b)

(SEQ ID NO: 195)
10 ATGCATATTGAGACTGTTATTGATTTCAAAGAATTAGGAAAAAGATATCGT
TTTAAAAATCCTACAAAAGAATTAATAGCTGATACTTTAGAACAAGTCTTA
GAAGTGATAAAAGAAGTTGATTATTATCAATCTCAAAATTATTATGTTGTT
GGTTATTTATCTTATGAAGCATCTGCTGCTTTTGATTACATTTTAAAGTTT
CTCAACAGAAGTTGGCTGGAGAACATCTAGCTTATTTTACAGTACATAAAG
ATTGTGAGAACGAAGCTTTTCCTTTAAGTTATGAAAATGTTAGATTAGCAG
ATAATTGGACTGCTAATGTTTCTGAGCAAGAATATCAAGAGGCAATTGCTA
15 ATATTAAAGGACAAATTAGACAAGGAAATACTTATCAAGTAAATTATACA
CTAGAGCTTAGCCAACAATTATGCTCGGATCC

(SEQ ID NO: 196)
20 MHIETVIDFKELGKRYRFKNPTKELIADTLEQVLEVIKEVDYYQSQNYVVG
LSYEASAAFD SHFKVSQQKLAGEHLAYFTVHKDCENEAFPLSYENVRLADNW
TANVSEQEYQEALANIKGQIRQGNTYQVNYTLELSQQLCSD

Sequence description:

25 A] Length: 440 bp - 146 aa (partial gene sequence)
B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-103 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-103 gene sequence.
30 Shine Dalgarno sequence present upstream of
ATG start codon, No apparent leader peptide sequence

ID-176

35

Clone 2-18/22b(b) (ID-104b)

(SEQ ID NO: 197)
40 GTGAATAATATGTTTTATCTCAAAATAGCCTGGCATAATTAAAACATTCT
ATAGACCAGTACATACCATTCTTAGCCAGTTTATTACTTTATTCATTGA
CTTGTTCTACGCTACTAATCTTAATGAGTGCTGTTGGAAGAGATATGGGGA
CAGCGGCAACGGTTCTTTTTCTTGGAGTGATTGTTTTGTCAATCTTTGCGGT
AGTCATGGAACATTATAGCTACAATATCTTGATGAAACAGCGTAGTAGTG

AATTTGGACTGTATAACATTTTGGGGATGAATAAACGTCAAGTTGCGCGTG
 TAGCTAGTCTAGAGCTGTTTATTATTTATATATTTCTTATTTCTATAGGAAG
 TCTGTTTAGTGCTTTTTTTTGCTAAATTTATTTATTTAATTTTTGTCAACATTA
 TTAACATCATGCACTAAATCTTAGTTTAAGTTTATGGCCATTTATTATTG
 5 TATCGTTATATTTACAGGTATTTTCTGACTTTAGAAAGTTCCAGTTATTCGA
 CATGTTTCATTTATCATCCCCATTAAGTCTTTTATAGAAAGAAACAACAGGGA
 GAAAAAGAACCAAAAAGGTAATCTTATACTTGCAATTTTATAGCGTTAGTAGCT
 ATCGCCATCGCTTATACAATGGCTCTTACTTCAGGTAAAGCACCTGCATTA
 GCTGTTATCTATCGTTTCTTCTTTGCAGTACTTTTAGTAATTGCTGGTACTT
 10 ATCTTTTTTATATTAGTTTTATGACATGGTACTTAAAAAGGTTGCGTCAAAA
 CAAGCATTATTATTATAAATCTGAGCATTTTGTATCAACTTCGCAAATGAT
 TTTTCGAATGAAGCAAAATGCAGTAGGGTTAGCAAGTATCACTTTATTAGC
 TGTTATGGCTCTAGTTACTATTGCTACAACAGTCTCACTCTATTCAAATACA
 CAAAATGTTGTTACCGGACTATTTCCAAAATCAGTAAGTTTATCAATAGAT
 15 AATTCAAAAGGTGACGCGAAAAATATATTTGAAGAAAAGATTTTGAAGAA
 ACTAGGTAAGTCATCTAAGGAAGCTATCACTTATAATCAGACAATGATTTC
 GATGCCAGTTAGTCAATCAAGTGAAGTAAATATCACATCTA

20 (SEQ ID NO: 198)
 MNNMFYLNKIAWHNLKHSIDQYIPFLLASLLLYSLTCSTLLILMSAVGRDMGTA
 ATVLFLGVIVLSIFAVVMEHYSYNILMKQRSSEFGLYNILGMNKRQVARVASL
 ELFIYIFLISIGSLFSAFFAKFIYLFVNINYNHALNLSLSLWPFIIICIVIFTGIFLTLE
 VPVIRHVHLSSPLSLFRKKQQGEKEPKGNLILAILALVAIAIAYTMALTSGKAP
 25 ALAVIYRFFFAVLLVIAGTYLFYISFMTWYLNKRLRQNKHYYYKSEHFVSTSQM
 IFRMKQNAVGLASITLLAVMALVTIATTVSLYSNTQNVVTGLFPKSVSLSIDNS
 KGDAKNIFEEKILKKLGKSSKEAITYNQTMISMPVSQSDDLISHL

30

Sequence description:

- 35 A] Length: 1119 bp - 373 aa (partial gene sequence)
 B] This gene sequence was not identified using the LEEP system. It was
 identified upstream of the ID-104 gene which was identified by LEEP, during
 cloning and sequence analysis of the full-length ID-104 gene sequence.
 Possible Shine Dalgarno sequence present
 40 upstream of a GTG start codon. Possesses a potential
 leader peptide sequence

ID-177

Clone 2-5b (ID-112b)

(SEQ ID NO: 199)
5 ATGGTTGAGCCAATTATTTCAATACAAGGACTTCATAAAAGTTTTTGGGAAA
AATGAGGTTTTAAAAGGCATTGACTTGGATATTCATCAAGGAGAAGTGGT
GGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTTTAAGAACAAT
GAATCTCTTGGAAGTACCAACAAAGGGGAACAGTGACTTTTGAAGGGATTG
ATATAACAGACAAAAAGAATGATATTTTTTAAAATGCGCGAAAAAATGGGC
10 ATGGTTTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAAT
ATTACTTTATCACCTATTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAG
ACAAAAGCATACGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAA
TGCTTATCCAGCAAGCTTATCTGGAGGACAACAACACGGATTGCTATTGC
AAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTTGATGAACCTACTTCA
15 GCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTA
GCTAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCA
CGTGAAGTAGCGGATCGTGTCATTATTATGGATGCAGGGATTATTGTTGAG
CAAGGGACCCCTAAGAAAGTATTTGAGCAGACAAAAGAAATCCGCACAAG
AGACTTCTTAAGTAAAGTATTATAA

20

(SEQ ID NO: 200)
MVEPIISIQGLHKSFGKNEVLKGIDLDIHQGEVVVIIGPSGSGKSTFLRTMNLLE
VPTKGTVTFEGIDITDKKNDIFKMREKMGMVFQQFNLFPMNTVLENITLSPIKT
KGLSKLDAQTKAYELLEKVGLKEKANAYPASLSGGQQQRIAIARGLAMNPDV
25 LLFDEPTSALDPEMVGEVLTVMQDLAKSGMTMVIVTHEMGFAREVADRVIF
MDAGIIVEQGTPKKVFEQTKEIRTRDFLSKVL*

30 Sequence description:

A] Length: 735 bp - 244 aa (full length gene)
B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-112 gene which was identified by LEEP,
35 during cloning and sequence analysis of the full-length ID-112 gene sequence.
Shine-Dalgarno sequence precedes the 'ATG'
start codon. No obvious leader peptide

40 ID-178

Clone 2-5c (ID-112c)

(SEQ ID NO: 201)

ATGTCTCA^sTATCAAGAGTGGTTAGAAAACGACTCACTCGGTAAAGATATT
AAGTCAGATTTAGAAGCTATTAAGGAGATGAATCTGAAATTCAGGATCG
TTTTTACAAAACATTAGAATTTGGAACGGCGGGATTGAGAGGTAAACTTG
GAGCAGGAACCAATCGTATGAATACTTATATGGTGGGGAAAGCAGCACAA
5 GCATTAGCTAATCGATTATTGATCATGGCCCTGAAGCTATTGCACGTGGAA
TTGCAGTTAGTTATGATGTCCCGTTATCAATCTAAGGAATTTGCAGAATTA
ACTTGGTCCATTATGGCAGCAAATGGTATTAAAGCCTTATATTTA

(SEQ ID NO: 202)

10 MSHMNYKEIYQEWLEND^sSLGKDIKSDLEAIKGDSEIQDRFYKTLEFGTAGLR
GKLGAGTNRMNTYMGVKAQAALANRLLIMALKLLHVELQLVMMSRYQSKE
FAELTWSIMAANGIKALYL

Sequence description:

15 A] Length: 366 bp - 122 aa (partial gene sequence)
B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-112 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-112 gene sequence.
20 Shine-Dalgarno sequence preceded the 'ATG'
start codon. No obvious potential leader peptide sequence.

ID-179

25 Clone 2-5d (ID-112d)

(SEQ ID NO: 203)

ATGCAACCTGTAAAAGTCGATGAACCTTCTGTTGAAGAAACCATTACTATT
TTGAAAGGTATCCAAAAAATAACGAAGATTATCATCACGTAAAATATAA
30 TAATGATGCCATAGAAGCAGCTGCAGTACTATCTAATCGTTATATCCAAGA
CCGCTTTTTACCTGATAAAGCAATAGACTTATTAGATGAAGCTGGTTCTAA
AATGAACCTAACACTAAATTTTGTGATCCAAAAGAAATTGATCAACGTCT
CATTGAAGCAGAAAATTTAAAAGCGCAAGCGACTCGTGAAGAAGATTACG
AACGTGCAGCTTACTTCCGTGACCAGATTGCAAAATATAAAGAAATGCAG
35 CAACAAAAGGTCGACGATCAAGATACACCTATTATTACCGAAAAACAAT
TGAGCACATCATTGAAGAAAAAACGAATATCCCTGTTGGTGATTTAAAAG
AAAAAGAACAATCTCAATTAATTAATCTCGCAGATGACTTGAAACAGCAT
GTGATCGGCCAGGATGACGCTGTCATTAAGATTGCAAAAGCTATTCGTCGT
AATCGAGTTGGTCTTGGTAGCCCAAACCGTCCTATTGGFTCCTTTTTATTG
40 TAGGACCAACCGGTGTTGGTAAAACTGAACTTTCTAAACAAC^sTAGCAATTG
AGCTCTTTGGTTCAGCTGATAGTATGATTCGTTTTGATATGTCAGAGTACAT
GGAAAAGCATGCTGTTGCTAAATTAGTCGGAGCGCCTCCAGGATACGTGG
GATACGAGGAAGCTGGACAAC^sTAAGTAAAAGGTTTCGTCGAAATCCTTAC
TCGCTCATCCTTCTAGATGAAATTGAAAAAGCTCATCCCGATGTCATGCAT

5 ATGTTCTTGCAGGTCCTTGATGACGGTCGATTAACAGATGGACAAGGAAG
AACTGTTAGTTTTAAAGATACCATTATCATCATGACCTCAAATGCTGGTTC
TGGTAAAACTGAAGCAAGTGTTGGCTTTGGTGCCTCACGAGAAGGTAGGA
CGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCAT
GCAAGC

(SEQ ID NO: 204)
10 MQPVKVDEPSVEETITILKGIQKKYEDYHHVKYNNDAIEAAVLSNRYIQDRF
LPDKAIDLLDEAGSKMNLTLNFVDPKEIDQRLIEAENLKAQATREEDYERAAY
FRDQIAKYKEMQQQKVDDQDTPIITEKTIEHIEEKTNPVGDLEKEQSQLINL
ADDLKQHVIGQDDAVIKIAKAIRNRVGLGSPNRPIGSFLFVGPTGVGKTELSK
QLAIELFGSADSMIRFDMSEYMEKHAVAKLVGAPPGYVGYEEAGQLTEKVRR
NPYSLILLDEIEKAHPDVMHMFLLQVLDDGRLTDGQGRTVSFKDTHIMTSNAGS
GKTEASVGFGASREGRTNSSSVPGDPLESTCRHAS

15

Sequence description:

20

A] Length: 1070 bp ÷ 356 aa (Partial gene sequence)

25 B] This gene sequence was not identified using the LEEP system. It was
identified upstream of the ID-112 gene which was identified by LEEP, during
cloning and sequence analysis of the full-length ID-112 gene sequence. Shine-
Dalgarno sequence preceded the 'ATG'
start codon. No obvious potential leader peptide
sequence.

ID-180

30

Clone 2-7b (ID-113b)

(SEQ ID NO: 205)
35 ATGAGAGGGAAGGTTATTTACGGCACAACCCTTATAGGTCTTTTTCTATTC
TTATTTTTCTATTTTTGGATTCCCTAAGCATCACATCGAGAGAATACATCATC
ATCGTATAAAGCAGGTAGATGCGAAGAGTGATTTAACAGGATTTAAAACC
CATTTGCCCATTCAGCATTGATACAAAGCAACAAGTTATTCCTCTTGTT
ACAAAAGAAGGCGGAAAATATGTCAAAGCTAGGGATAATATTAATGTTGA
TATCGAATTACGGGATTCTCCAAGTAGATCACATCATTATCAGAAAAGCC
40 GAGAATTAGGACAAAAGGGTTAATATCATATAGAGGAAATTCCTCTCGTT
ACTTTGATAAGAAGTCATTGAAAGTTAAGTTTGTTACTAATAAGTTAAAGG
AAAAGAAGCATCGATTAGCAGGAATGCCTAAAGAATCGGAGTGGGTATTG
CATGGTCCCTTTCTAGACAGAACATTATTAAGAAATTATCTGAGTTATAAT

ATTGCTGGTGAGATTATGCCTATGCCCCAAACGTTTCGCTACTGTGAGTTAT
TTGTCAATGGTGAGTATCAGGGAG

(SEQ ID NO: 206)

5 MRGKVIYGTTLIGLFLFLFFYFWIPKHHIERIHHHRIKQVDAKSDLTGFKTHLP
SIDTKQQVIPLVTKEGGKYVKARDNINVDIELRDSPSRSHHLSEKPRIRTKGLIS
YRGNSSRYFDKKSLKVKFVTNKLKEKKHRLAGMPKESEWVLHGPFLLDRTLLR
NYLSYNIAGEIMPMPQTFATVSYLSMVSIRE

10

Sequence description:

A) Length: 582 bp - 194 aa (Partial gene sequence)

15 B) This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-113 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-113 gene sequence.
ATG start codon is preceded by a Shine-
Dalgarno sequence-Possesses a potential leader peptide
sequence. C-terminus to be determined.

20

ID-181

25

Clone 2-17b (ID-117b)

(SEQ ID NO: 207)

30 CTTACATTTTATTGATCACTATCTGACAAATGTTAATCAAACAGCAGTTCT
TATTTTAGTGGGATATTATTCAATGTATGTCTTGCAGACCTTAATTCAATAT
TTTGGGAATCTCTTTTTTTCGCGCGTGTTTCTTATAGTATTGTTAGAGATATTC
GTAGAGATGCTTTTGCTAATATGGAAAGGCTAGGCATGTCTTATTTTGATA
GGACACCGGCAGGATCTATTGTGTACGTATTACTAATGATACTGAAGCAA
TATCTGATATGTTTTTCGGGTATTTTATCAAGTTTTATCTCGGCGATATTTAT
35 TTTTACAGTTACTCTGTACACTATGTTGATGCTAGACATTAACTAACAGG
ACTCGTCGCTCTTTTGTTACCTGTTATCTTTATATTAGTGAATGTCTATCGG
AAAAAATCAGTCACTGTCATTGCTAAAACGAGAAGTTTACTTAGTGATATC
AACAGTAAATTATCAGAAAGTATTGAAGGAATTC

(SEQ ID NO: 208)

40 SHFIDHYLTNVNQTAVLILVGYYSMYVLQTLIQYFGNLFARVSYSIVRDIRRD
AFANMERLGMSYFDRTPAGSIVSRITNDTEAISDMFSGILSSFISAIFIFTVTLYT
MLMLDIKLTGLVALLLPVIFILVNVYRKKSVTVIKTRSLSDINSKLSIESIEGI

Sequence description:

- 5 A] Length: 498 bp - 165 aa (Partial gene sequence)
 B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-117 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-117 gene sequence. N- and C-termini have yet to be determined

10

ID-182

15

Clone 3-8b (ID-120b)

(SEQ ID NO: 209)
20 ATGTACCATATTGAATTAAAAAAGGAAGCTTTACTACCAAGAGAACGCCT
AGTTGATTTAGGCGCAGATAGATTGAGTAATCAGGAGTTATTAGCCATTCT
CTTACGTACAGGTATTAAAGAAAAACCTGTTCTTGAAATTTCAACGCAAAT
TTAGAAAAACATAAGCAGTTTAGCAGATTTTGGTCAATTATCCTTACAGGA
GTTGCAATCCATTAAAGGAATCGGTCAGGTAAATCCGTCGAAATAAAAG
CTATGCTAGAACTAGCAAAACGGATTCACAAAGCTGAATATGATCGTAAA
25 GAGCAAATTTTAAGTAGTGAACAATTAGCGAGGAAAATGATGCTCGAATT
AGGGGATAAAAAACAAGAACATTTAGTAGCTATTTATATGGATACACAAA
ATCGTATTATCGAACAGAGAACTATTTTTATTGGTACTGTACGTCGTTCAG
TAGCAGAGCCAAGAGAAATTCTACATTATGCTTGTA AAAACATGGCAACT
TCTTTGATTATTATACATAATCATCCCTCAGGTTCTCCAAATCCCAGTGAAA
30 GTGATTTAAGTTTCACTAAAAAAATAAAACGATCATGTGATCATCTGGGAA
TTGTCTGCCTAGATCACATCATCGTTGGAAAAAATAAATATTATAGTTTTC
GAGAAGAAGCAGATATTTTATAA
(SEQ ID NO: 210)
35 MYHIELKKEALLPRERLVDLGADRLSNQELLAILLRTGIKEKPVLEISTQILENI
SSLADFGQLSLQELQSIKGIGQVKSVEIKAMLELAKRIHKA EYDRKEQILSSEQ
LARKMMLELGDKKQEHLVAIYMDTQNR IIEQRTIFIGTVRRSVAEPREILHYAC
KNMATSLIIHNHPSGSPNPSESDLSFTKKIKRSCDHLGIVCLDHIIVGKNKYYSF
REEADIL*

40

Sequence description:

- A] Length: 681 bp - 227 aa (full-length gene)
B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-120 gene which was identified by LEEP,

during cloning and sequence analysis of the full-length ID-120 gene sequence.
ATG start codon is preceded by an typical
Shine-Dalgarno sequence. No obvious leader
peptide sequence

5

ID-183

10

Clone 3-11b (ID-121b)

(SEQ ID NO: 211)
15 TGGTTAAAAGTAGTGATAGCTTGTATTCCATCTATTTTAATTGCTTTACCAT
TTGATAATTGGTTTGAAGCTCATTTTAATTTTCATGATTCCGATTGCAATAGC
CCTAATCTTTTATGGTTTTGTCTTCATATGGGGTTGAAAAACGTAATGCACAC
CTCAAACCACAGGTAACCGAATTGGCAAGTATGTCCTTACAAGACAGCTTTC
TTGATTGGATGTTTCCAGGTTCTCAGTATTGTTCCGGGAACCAGTCGTTCTG
GAGCTACTATTTTAGGAGCAATTATTATTGGAAGTAGTCGTTCCGGTCGCTG
20 CTGACTTTACTTTCTTCCTTGCCATCCCAACTATGTTTGGTTATAGTGGACT
TAAGGCGGTAAATATTTTTTAGATGGTAACGTCTTGAGTTTAGACCAATC
TTTAATACTTTTAGTAGCAAGTCTGACAGCTTTCGTAGTTAGTTTATATGTT
ATTCGTTTCTTGACAGACTATGTCAAACGACACGATTTACCATCTTTGGT
AAGTATCGTATAGTCTTAGGAAGTTTACTCATCCTCTACTGGTTAGTTGTTC
25 ATTTATTCTAA

(SEQ ID NO: 212)
30 WLKVVIACIPSILIALPFDNWFEAHFNFMIPIALIFYG FVFIWVEKRNAHLKP
QVTELASMSYKTAFLIGCFQVLSIVPGTSRSGATILGAIIGTSRVAADFTFFLA
IPTMFGYSGLKAVKYFLDGNVLSLDQSLILLVASLTA FVVS LYVIRFLTDYVKR
HDFTIFGKYRIVLGSLLILYWL VVHLF*

Sequence description:

35 A] Length: 579 bp - 193 aa (partial sequence)
B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-68 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-68 gene sequence
described in WO 00/06736. N-terminus has yet to be determined:

40

ID-184

Clone 3-11c (ID-121c)

(SEQ ID NO: 213)

ATGGAAATGAAACAAATCAGTGAAACAACACTGAAAATTACAATTAGTAT
GGAAGATTTAGAAAGATCGTGGTATGGAGCTGAAAGATTTTCCTAATCCCTCA
5 GGAGAAGACTGAGGAATTTTTCTATTCTGTCATGGATGAATTAGACTTGCC
AGAAAACCTTTAAAAATAGTGGTATGTTAAGTTTTTCGAGTAACACCTAAAA
AAGATCGCATTGATGTTTTTGTACAAAGTCTGAATTAAGTAAAGATTTAA
ATTTAGAAGAATTAGCAGATTTGGGTGACATTTCAAAAATGTCTCCAGAAG
ACTTTTTTAAACCTTGGAACAATCGATGTTGGAAAAAGGGGATACGGAT
10 GCCCATGCCAAATTAGCAGAAATTGAAAATATGATGGATAAAGCAACTCA
AGAAGTAGTTGAGGAAAATGTTTCTGAAGAACAACCTGAAAAGGAAGTAG
AAACGATTGGATATGTTCACTATGTCTTTGATTTTGATAATATTGAAGCTGT
AGTTCGATTTTCACAAACGATTGATTTTCCAATAGAAGCTT

(SEQ ID NO: 214)

MEMKQISETTLKITISMEDLEDRGMELKDFLIPQEKTEEFFYSVMDELDLPENF
KNSGMLSFRVTPKKDRIDVFVTKSELSKDLNLEELADLGDISKMSPEDFFKTLE
QSMLEKGDTHAKLAEIENMMDKATQEVVEENVSEEQPEKEVETIGYVHY
VFDFDNIEAVVRFSQTIDFPIEA

20

Sequence description:

A] Length: 547 bp - 182 aa (Partial sequence)
B] This gene sequence was not identified using the LEEP system. It was
25 identified downstream of the ID-68 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-68 gene sequence.
ATG start codon is preceded by an typical
Shine-Dalgarno sequence. No obvious potential
leader peptide
30 sequence

35 ID-185

Clone 3-16b (ID-122b)

(SEQ ID NO: 215)

GGAAACCAACGGCCAGTACAATCGTCAAGGGTAGATTATCCTAAACGTAG
TCGTGCCAAGATTGTAGAAGTTTATTTTAGACAAGCTTCTACTACTGATTA
TTCTGGTGTTTACAAAGGTTACTATATTGACTTTGAAGCCAAAGAAACCCG
GCAGAAAACCTGCTATGCCTATGAAAAATTTTCATGCTCACCAAATAGAGC
ACATGGCAAATGTATTACAGCAAAAAGGGATTTGCTTTGTCTTGCTTCATT

TTTCCACACTTAAGGAAACCTATCTACTCCCTGCTAATGAGTTAATTTTCATT
TTATCAGATTGATAAAGGCAATAAATCAATGCCTATTGATTATATCAGAAA
AAATGGATTTTTCGTAAAGGAGAGTGCCTTTCCTCAAGTCCCTTACTTAGA
TATTATTGAAGAAAAATTATTAGGCGGTGATTACAATTAA

5

(SEQ ID NO: 216)

GNQRPVQSSRVDYPKRSRAKIVEVYFRQASTTDYSGVYKGYIDFEAKETRQ
KTAMPMKNFHAHQIEHMANVLQQKGICFVLLHFSTLKETYLLPANELISFYQI
DKGNKSMPIDYIRKNGFFVKESAFPQVPYLDIIEKLLGGDYN*

10

Sequence description:

A) Length: 447 bp - 149 aa (partial sequence)

15

B) This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-122 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-122 gene sequence. N-terminus has yet to be determined

20

ID-186

Clone 3-17b (ID-123b)

25

(SEQ ID NO: 217)

GGATCCTAAAAACGCTAAGGTTTATCAAAAAAATGCTGATCAATTTAGTG
ACAAGGCAATGGCTATTGCAGAGAAGTATAAGCCAAAATTTAAAGCTGCA
AAGTCTAAATACTTTGTGACTTCACATACAGCATTCTCATACTTAGCTAAG
30 CGATACGGATTGACTCAGTTAGGTATTGCAGGTGTCTCAACCGAGCAAGA
ACCTAGTGCTAAAAAATTAGCCGAAATTCAGGAGTTTGTGAAAACATATA
AGGTTAAGACTATTTTTTGTGGAAGAAGGAGTCTCACCTAAATTAGCTCAAG
CAGTAGCTTCAGCTACTCGAGTTAAAATTGCAAGTTTAAGTCCTTTAGAAG
CAGTTCCCAAAAACAATAAAGATTACTTAGAAAATTTGGAAACTAATCTTA
35 AGGTACTTGTCAAATCGTTAAATCAATAG

35

(SEQ ID NO: 218)

DPKNAKVYQKNADQFSDKAMAIAEKYKPKFKAASKYFVTSHTAFSYLAKR
YGLTQLGIAGVSTEQEPSAKKLAEIQEFVKTYKVKTIFVEEGVSPKLAQAVAS
40 ATRVKIASLSPLEAVPKNNKDYLENLETNLKVLVKS LNQ*

40

Sequence description:

A] Length: 433 bp - 144 aa (partial sequence)
B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-123 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-123 gene sequence. N-terminus has yet to be determined

ID-187

Clone 3-46/47 (ID-130b)

(SEQ ID NO: 219)
15 ATGAAAAAAGTCATCGATTAAAAAACTACAAAAAGCATACGCCTCAGA
AACTGTTTTAAATAATATTAATTTGGAGGTGTTTAAAGGAGAAATAATTGG
ATTAATAGGACCCTCTGGAGCAGGGAAATCTACCTTGATTAAAACTATGCT
TGGCATGGAAAAAGCAGATAAGGGAACAGCTCTTGTTCTTGATACTCAAA
TGCCAGATCGTAATATTTTAAATCAAATTGGCTATATGGCTCAATCTGATG
20 CCTTACACGAGTCTTTAACTGGCTTAGAAAAATTTATTATTCTTTGGAAAAA
TGAAAGGTATTCAAAAAAAGTGAATTAACAGCAGATAACTCATATTTCT
AAAGTAGTAGATCTAGAAAACCAACTTGATAAATTTGTCTCAGGTTACTCA
GAAGGTATGAAAAGACGGCTTTCTCTAGCCATCGCCCTACTTGGAAACCCC
ACAGTTTTAATCCTAGATGAACCTACCGTTGGAATTGATCCATCCTTGAGG
25 AGAAAAATCTGGCAAGAGCTAATTAATATTAAGGATGAAGGACGTTCTAT
CTTTATTACAACCCACGTTATGGATGAAGCAGAATTAACAAGTAAGGTTGC
ACTACTATTACGTGGAAACATTATTGCCTTTGATACTCCATTACATTTAAA
AAAACAATTTAATGTGAGTACTATTGAGGAAGTTTTCTTAAAAGCTGAAGG
AGAATAA
30 (SEQ ID NO: 220)
MKKVIDLKKLQKAYASETVLNNINLEVFKGEIIGLIGPSGAGKSTLIKTMLGME
KADKGTALVLDTQMPDRNILNQIGYMAQSDALHESLTGLENLLFFGKMKGIO
KTELKQQITHISKVVDLENQLDKFVSGYSEGMKRRLSLAIALLGNTPTVLILDEP
TVGIDPSLRRKIWQELINIKDEGRSIFITTHVMDEAELTSKVALLLRGNIIAFDTP
35 LHLKKQFNVSTIEEVFLKAEGE*

Sequence description:

40 A] Length: 717 bp - 239 aa (Possible full-length sequence)
B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-130 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-130 gene sequence. ATG start codon is preceded by a possible

Shine-Dalgarno. No obvious potential leader
peptide sequence

5

ID-188

Clone 3-83b (ID-144b)

10

(SEQ ID NO: 221)
ATGGTACAAATGATACATGATATGATTAACAATTGAGCATTTTGCTGAG
ACACAAGCTGATTTTCCAGTGTATGATATTTTAGGGGAAGTCCATACTTAT
GGACAACCTTAAAGTAGACTCTGACTCTCTAGCTGCTCATATTGATAGCCTA
GGCCTTGTTGAAAAATCACCTGTCTTAGTATTCGGTGGTCAAGAATATGAA
15 ATGTTGGCGACATTTGTTGCTTTAACAAAGTCAGGGCATGCTTATATACCG
GTTGACCAACACTCTGCTTTGGATAGAATACAGGCTATTATGACAGTTGCT
CAACCAAGCCTTATCATTTCAATTGGTGAATTTCTCTTGAAGTTGATAAT
GTCCCAATCCTAGACGTTTCTCAAGTTTCAGCTATTTTGAAGAAAAGACT
CCTTATGAGGTAACACATTCTGTAAAGGTGATGATAATTACTATATTATT
20 TTCACTTCAGGGACTACTGGTTTACCAAAGGTGTGCAAATTTACATGAC
AATTTATTGAGCTTTACAAATTGGATGATTTCTGATGATGAGTTTTTCAGTTC
CTGAAAGACCGCAAATGTTGGCTCAACCC

20

(SEQ ID NO: 222)
MVQMIHDMIKTIEHFAETQADFPVYDILGEVHTYGQLKVDSDSLAAHIDSLGL
25 VEKSPVLVFGGQEYEMLATFVALTKSGHAYIPVDQHSALDRIQAIMTVAQPSL
IISIGEFPLEVDNVPILDVSQVSAIFEEKTPYEVTHSVKGDDNYYIIFTSGTTGLP
KGVQISHDNLLSFTNWMISDDEFSVPERPQMLAQP

25

30

Sequence description:

35

A] Length: 592 bp - 197 aa (partial sequence)
B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-144 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-144 gene sequence.
Putative ATG start codon is preceded by a
typical Shine-Dalgarno sequence. No obvious
leader peptide sequence

40

This orf is not in frame with nuc

ID-189

Clone 3-86b (ID-145b)

5 (SEQ ID NO: 223)
ATGGAAAATCATCGTTATGAAGATGAAGGTAAATTCCAGCGTAAGATGAC
CAGTCGTCATCTCTTTATGTTATCGCTAGGTGGTGTATCGGGACTGGGCTT
TTCTTGAGTTCAGGTTATACCATTCACAGGCTGGTCCGCTTGGAGCTGTG
CTGTCTTATTTGATTGGTGCCGTTGTGGTTTATTTGGTCATGCTATCACTTG
10 GGAATTGGCGGTTGCCATGCCGGTGACGGGGGTCATTCCACACTTATGCCA
CTAAGTTTATCAGTCCTGGAACAGGTTTTACTGTTGCTTGGCTATATTGGAT
TTGTTGGACGGTCGCCTTGGGGACTGAATTTTTAGGTGCTGCCATGCTGAT
GCAGCGCTGGTTCCCAAATGTGCCGGCTTGGGCATTTGCTTCCTTTTTTGGC
CTTGTGATTTTTTGGTTTAAATGCTCTTAGCGTACGCTTTTTTGCAGAAGCAG
15 AGTCTTTCTTCTCAAGTATTAAGGTTATTGCTATCATTATCTTTATTATCTTG
GGCTTAGGTGCTATGTTTGGTCTAGTTTCCTTTGAAGGTCAGCACAAGGCT
ATTCTCTTCACTCATCTGACTGCCAATGGTGCCTTTCCAAATGGTATCGTTG
CAGTTGTCTCAGTCATGTTGGCTGTTAACTATGCCTTCTCTGGTACTGAGTT
AATTGGTATTGCGGCTGGTGAAACGGATAATCCCAAAGAAGCTGTACCAA
20 GGGCTATTAACGACAATCGGTCGCTTGGTTGTTTTCTTTGTACTGACAA
TTGTTGTCCTAGCTTCGCTATTGCCAATGAAAGAGGCAGGCGTATCCACAG
CACCATTTCGTTGATGTCTTTGACAAGATGGGAATCCCTTTTACGGCGGATA
TCATGAACTTCGTTATCTTGACAGCCATCCTGTCTGCTGGTAACTCAGGTCT
CTACGCATCAAGCCGTATGCTCTGGTCCCTTGCCAATGAAGGTATGTTGTC
25 AAAATCTGTTGTGAAAATCAATAAACACGGTGTCCCAATGCGTGCTCTTCT
CTTGTCAATGGCAGGAGCAGTGCTGTCGCTCTTTTCAAGTATTTACGCTGC
AGACACAGTTTATCTAGCCTTGGTTTCAATCGCGGGGCTTTGCTGTTGTTGTC
GTATGGCTAGCCATTCCAGTCGCACAAATCAATTTCCGCAAGGAATTC
30 (SEQ ID NO: 224)
MENHRYEDEGKFQRKMTSRHLFMLS LGGVIGTGLFLSSGYTIAQAGPLGAVL
SYLIGAVVVYLVMLS LGELAVAMPVTGSFHTYATKFISPGTGFTVAWLYWIC
WTVALGTEFLGAAMLMQRWFPNVPWAFASFFALVIFGLNALSVRFFAEAES
FFSSIKVIAIIIFILGLGAMFGLVSFEGQHKAILFTHLTANGAFPNGIVAVVSVM
LAVNYAFSGTELIGIAAGETDNPKEAVPRAIKTTIGRLVVFFVLTIVVLASLLPM
35 KEAGVSTAPFVDVFDKMGIPFTADIMNFVILTAILSAGNSGLYASSRMLWSLA
NEGMLSKSVVKINKHGVPMRALLSMAGAVLSLFSSIYAADTVYLALVSIAGF
AVVVVWLAIPVAQINFRKEF

40

Sequence description:

A) Length: 1126 bp - 393 aa (partial gene

sequence)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-145 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-145 gene sequence. Putative ATG start codon is preceded by a typical Shine-Dalgarno sequence. Possesses a possible leader peptide sequence.

5

10

ID-190

15

Clone 3-94b

20

25

30

35

(SEQ ID NO: 225)
TCAGAAAATGCAGAGGCAGCAACGGTTGCCACAACTTGGTTACCAAAGG
AGCTAATGTCATTATCGGACCAGCAACATCGGGTGCAGCTGCATCTTCAAC
TCCAAAAGTAAATGCAGCAGCAGTTCCAATGATTGCACCTGCTGCGACAC
AAGACAATTTAGTCTATGGTTCTGATGGAAAAACCTTAAATCAGTATTTCT
TCCGAGCTACTTTTGTGCGATAATTATCAAGGAAAGCTATTGTCTCAGTATG
CTACAGACAACCTTAAAGCTAAAAAAGTTGTTCTATTTTATGATAATTCAT
CAGATTACTCAAAGGGGGTAGCAAAATCATTTAAGGAAAGTTATAGTGGA
AAAATTGTTGATAGTATGACATTCTCCGCTGGTGATACTGATTTCCAAGCG
TCATTGACTAAGTTGAAAGGGGAAAGAATATGATGCTATTGTGATGCCAGG
TTACTATACCGAGACAGGATTAATAGTTAAGCAAGCGCGTGATTTAGGTAT
CTCTAAACCGGTTCTTGGGCCTGATGGTTTTGATAGTCCGAAATTTGTGCA
ATCGGCAACACCTGTAGGAGCTTCAAACGTTTATTATTGACAGGTTTCAC
TACACAAGGATCAACCAAAGCTAAAGCT

(SEQ ID NO: 226)
SENAEAATVATNLVTKGANVIIGPATSGAAASSTPKVNAAAVPMIAPAATQD
NLVYGSDGKTLNQYFFRATFVDNYQGKLLSQYATDNLKAKKVVLFDNSSD
YSKGVAKSFKESYSGKIVDSMTFSAGDTDFQASLTKLKGKEYDAIVMPGYT
ETGLIVKQARDLGISKPVLGPDGFDSPKFVQSATPVGASNVYYLTGFTTQGST
KAKA

40

Sequence description

A] Length: 637 bp - 231 aa (partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-149 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-149 gene sequence. N- and C-termini have yet to be determined

5

ID-191

10

Clone 2-c94b (ID-153b)

(SEQ ID NO: 227)
15 TTGGGACTTAAAGACCATGCTTTAGTCTATCCATTTTCATTATCTGGGGGGG
CAAAAGCAACGTGTCGCACTAGCTCGTGCGATGATGATTGATCCACAGATT
ATTGGTTATGATGAGCCAACACTAGCGCTCTTGATCCAGAGTTGCGTCAAGAA
GTAGAAAACTAATTTTACAAAATAGAGAAACAGGTATGACACAAATTGT
AGTAACACATGATCTTCAATTTGCTGAAAGTATATCTGATACGATTCTCAA
AATTAATCCTAAGTAG
20 (SEQ ID NO: 228)
MGLKDHALVYPFSLSGGQKQRVALARAMMIDPQIIGYDEPTSAIDPELRQEV
EKLILQNRETGMTQIVVTHDLQFAESISDTILKINPK*

25

Sequence description

30 A] Length: 270 bp - 90 aa (partial sequence)
B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-153 gene which was identified by LEEP, during cloning and sequence analysis of the ID-153 gene sequence.
N-terminus has yet to be determined

35

ID-192

40 Clone 2-c1b (ID-155b)

(SEQ ID NO: 229)
ATGACTAATATCTCAGATGTTCCAAAAGCTATTAGAACACAGGCACAGTAT
GTTCTCTTGGGAATGAGAGTTATGGATCAGTCGGTATTACCGAAAACATAT
AATTCAAAAGAACCTTATTTGAAACCAGATATGATTTATATTCATGATAGA

AGACAAGAGACAATGCTTAAAATCACTCAAGAAATAGAAATGGAGCATTG
A

(SEQ ID NO: 230)

5 MTNISDVPKAIRTQAQYVLLGMRVMDQSVLPKTYNSKEPYLKPDMIYIHDRR
QETMLKITQEIEMEH*

Sequence description

10

A] Length: 204 bp - 68 aa (partial sequence)

15 B] This gene sequence was not identified using the LEEP system. It was
identified upstream of the ID-155 gene which was identified by LEEP, during
cloning and sequence analysis of the ID-155 gene sequence.

ATG start codon is preceded by a potential typical Shine-Dalgarno sequence.

Has a

typical leader peptide. N-terminus has yet to be
verified

20

ID-193

25 Clone 2-54altb (ID-172b)

(SEQ ID NO: 231)

AAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCTTGGGGGAATATAAATT
TGGATTTCATGACGATGTAAAGCCAATTTATTCTACGGGAAAAGGTCTAAA
TGAGGCTGTTATTCGTGAGTTATCTGCAGCTAAGGGTGAACCTGAGTGGAT
30 GTTGGACTTTCGTCTAAAATCCTTGGAAACGTTTAATAAAAATGCCGATGCA
GACCTGGGGAGCAGATTTATCAGATATTGATTTTGATGATATTATTTATTA
TCAAAAAGCATCTGATAAACCTGCGCGTGATTGGGATGATGTTCCAGAAA
AAATCAAAGAACTTTTGAAAGAATTGGGATTCCAGAAGCTGAAAGAGCC
TATCTTGCAGGAGCATCAGCACAAATATGAATCAGAAGTAGTTTATCACAAT
35 ATGAAAGAAGAATATGATAAGCTGGGTATTGTTTTTACGGATACTGACTCT
GCACTTAAAGAGTACCCAGAGCTATTCAAAAAATATTTTGCTAAACTTGTC
CCTCCAACAGATAATAAATTAGCTGCTCTGAACTCTGCTGTATGGTTCAGGT
GGAACATTTATTTATGTTTCCTAAAGGTGTTAAGGTGGATATTCCACTTCAA
ACTTACTTCCGTATTAATAATGAAAATACTGGACAATTTGAACGTAATCTC
40 ATTATTGTTGATGAGGGAGCAAGTGTTCACTATGTTGAAGGTGTACCGCC
CCAACCTATTCTTCAAATAGTTTACATGCAGCTATAGTTGAAATTTTGCAC
TTGATGGAGCTTATATGCGCTATACGACTATTCAAAATTGGTCCGATAATG
TCTATAATTTAGTGACAAAACGTGCTACCGCTAAAAAAGATGCAACAGTT
GAGTGGATAGATGGAAATCTAGGAGCTAAAACAACAATGAAATACCCATC

GGTTTACCTTGATGGTGAAGGAGCACGTGGCACGATGTTGTCTATTGCTTT
TGCAAACAAAGGACAACACCAAGATACGGGTGCAAAGATGATTCATAATG
CCCCCATACTAGTTCATCCATTGTCTCTAAATCAATTGCTAAGGGTGGGG
GAAAAGTTGATTATCGAGGTCAAGTGACATTTAATAAAGATTCCAAAAAA
5 TCAGTGTCACATATAGAATGTGACACCATATTGATGGATGATATTTCAAAA
TCAGATACCATACCGTTTAATGAGATTCATAATTCACAGGTTGCTTTAGAG
CATGAAGCAAAGGTGTCTAAGATTTCTGAAGAGCAACTGTACTACTTGATG
AGTCGAGGTTTATCTGAAGCTGAAGCAACAGAAATGATTGTTATGGGGTTT
GTTGAGCCCTTTACGAAAGAATTACCAATGGAATATGCGGTAGAGTTAAA
10 TCGTTTAATTTCTATGAAATGGAAGGTTTCAGTTGGTTAA

(SEQ ID NO: 232)
MHACRSTLEDLGEYKFGFHDDVKPIYSTGKGLNEAVIRELSAAKGEPEWMLD
FRLKSLETFNKMPMQTWGADLSIDFDDIIYYQKASDKPARDWDDVPEKIKE
15 TFERIGIPEAERAYLAGASAQYESEVVYHNMKEEYDKLGIVFTDSDSALKEYP
ELFKKYFAKLVPPTDNKLAALNSAVWSGGTFIYVPKGVKVDIPLQTYFRINNE
NTGQFERTLIIVDEGASVHYVEGCTAPTYSSNSLHAAIVEIFALDGA YMRYTTI
QNWSDNVYNLVTKRATAKKDATVEWIDGNLGAKTTMKYPSVYLDGEGARG
TMLSIAFANKGQHQDTGAKMIHNAPHTSSSIVSKSIKGGGKVDYRGQVTFN
20 KDSKKS VSHIECDTILMDDISKSDTIPFNEIHNSQVALEHEAKVSKISEEQLYYL
MSRGLSEAEATEMIVMGFVEPFTKELPMEYAVELNRLISYEMEGSVG*

Sequence description:

25 A] Length: 1411 bp - 469 aa (Possible full-length gene).
B] This gene sequence was not identified using the LEEP system. It was
identified downstream of the ID-72 gene which was identified by LEEP,
during cloning and sequence analysis of the full-length ID-72 gene sequence.
30 No obvious Shine Dalgarno sequence upstream of
TTG start codon (insufficient sequence data). N
terminus needs verification.

35

ID-194

Clone 3-1b (ID-81b)

(SEQ ID NO: 233)
40 ATGATAGAATTCTTTTCTAATATCAGAACAGAGATTCCGCAGATGCCTTTA
CTTATCCATAGTTTGATTTTATCTGTCTTACCTTTTCTGATGTGGCTGACTTT
GGTTAATAGAGATAAGCCTTTGTATAAACTATTTGGAGTATCCTTTTAGG
ACTTCAGTTAATTACGATTTATACTTGGTTTTTCTGGGCAAATTGCCTTTA

TCTGAAAGTCTTCCCCTTTACCATTGTCTGAATAGGCATGTTTGTCGGTCTCTTA

(SEQ ID NO: 234)

MIEFFSNIRTEIPQMPLLIHSLILSVLPFLMWLTLVNRDKPLYKTIWSILLGLQLI
TIYTWFFWAKLPLSESLPLYHCRIGMFVGLL

Sequence description

A) Length: 261 bp - 87 aa (partial gene sequence)

B) This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-81 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-81 gene sequence. Sequence Characteristics: Possesses a potential leader peptide sequence Orf is preceded by a potential Shine-Dalgarno sequence.

ID-195

Clone RS-55b

(SEQ ID NO: 235)

AAGCTTGTGCAAAGTATTAAAGAGATAGGATTAGCTAATGCGCATTATTATTA
GCTGTTGCTCCGACAGGGTCAATCAGTTATCTTTCTTCTTGTACTCCGAGCC
TTCAACCGGTTGTATCACCTGTCGAAGTACGCAAGGAAGGAGCACTGGGG
AGGGTTTATGTAGCTGCTTATAAGATTGATGCAGATAATTATGTCTACTAC
AAAAAAGGAGCTTATGAAGTGGGATCTGAGGCGATTATCAATATTGCAGC
TGCCGCTCAAAAACACATTGATCAAGCTATTTTCGTTAACGCTTTTCATGAC
AGATCAAGCAACTACGCGAGATTTAAATAAAGCCTATATTCAAGCATTTA
AACAAAAATGTGCCTCTATTTATTATGTACGAGTGAGACAGGACATCCTAG
AAGGTAGCGAGAGTTATGATGATATGCTGGATGATTTCACTTCATCGGACT
TAGAAGACTGTCAATCCTGCATGATTAA

(SEQ ID NO: 236)

→KLVQSIKEIGLANAHLAVAPTGSISYLSSCTPSLQPVVSPVEVRKEGALGRV
YVAAWKIDADNYVYYKKGAYEVGSEAINIAAAAQKHIDQAISLTLFMTDQAT
TRDLNKA YIQAFKQKCASIYYVRVRQDILEGSESYDDMLDDFTSSDLEDCQSC
MI*

Sequence description:

A] Length 486 bp - 162 aa (Partial sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-87 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-87 gene sequence. N-terminus to be determined.

ID-196

Clone RS-59(ID-90b)

(SEQ ID NO: 237)

GTGAGGACATATATTACAAACTTGAATGGACATTCAATCACTAGTACAGC
ACAAATAGCTCAAAACATGGTAACAGATATAGCAGTAAGCTTAGGTTTTC
GTGAGCTGGGAATACATTCTTATCCGATTGATACTGATTCTCCTGAGGAAA
TGAGTAAGCGTTTAGATGGAATCTGTTCCGGACTTAGAAAAAATGATATTG
TCATATTTTCAGACACCTACATGGAACACTACAACCTTTTGATGAAAAATTAT
TTCACAAATTAAAAATATTTGGTGTAAAGATTGTTATTTTTATACATGATGT
TGTACCGCTAATGTTTGATGGAAATTTTTATTTGATGGATAGAACTATAGC
TTATTATAATGAAGCAGATGTTTAATAGCCCCTAGTCAAGCAATGGTCGAT
AAGCTT

(SEQ ID NO: 238)

MRTYITNLNGHSITSTAQIAQNMVTDIAVSLGFRELGIHSYPIDTDSPEEMSKRL
DGICSGLRKNDIVIFQTPTWNTTTFDEKLFHKLKIFGVKIVIFIHDVVPLMFDGN
FYLMDRTIAYYNEADVLIAPSQAMVDKL

Sequence description:

A] Length: 414 bp - 138 aa(partial gene)

B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-90 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-90 gene sequence. No obvious signal peptide, but a possible Shine Dalgarno sequence is present upstream of ATG start codon. C-terminus has yet to be determined.

ID-197

Clone RS-59c (ID-90c)

(SEQ ID NO: 239)

5 CATGGAAATGAAGTTGATGATGTTATTAGAAGGGCATTGAATATAATCAC
CTTATCTTTGCTTTTGATAATACCTGTCATAACAGAGAGTTAGTATTAGATA
GCAATATCATTTCTCACACAACCTGTGAACAATTGATAAATTTAATGAAAA
ATTTATCAGGCTCCATTATGTATTTGCTAGAGCAACAAAGAGAACAAACA
AGTAATGAAACAAAAGAGCGTTATAAAGAAATATTAGGAGGGGTATGGAA
ATGCCTAA

(SEQ ID NO: 240)

10 HGNEVDDVIRRAFEYNHLIFAFDNTCHNRELVLDSNIISHTTCEQLINLMKNLS
GSIMYLLEQQREQTSNETKERYKEILGGYGNA*

15 Sequence description:

A] Length: 261 bp - 87 aa(partial gene sequence)

20 B] This gene sequence was not identified using the LEEP system. It was
identified upstream of the ID-90 gene which was identified by LEEP, during
cloning and sequence analysis of the full-length ID-90 gene sequence. N-
terminus has yet to be determined

ID-198

25

Clone RS-70b (ID-93b)

(SEQ ID NO: 241)

30 ACATTTTTATATTATGTATTTGAAGACGTAGCCACCCAGTCAAATATGACT
GGGAAGATTTTATAGTATGTCTAAAGAAGAGTTGTCATATTTACCCGTTATT
AAACTTTTAAAGAATCAAGGTGTATACAACGGCTTGATTGGTCTATTCCTC
CTTTATGGGTTATATATTTTACAGAAATCAAGAAATTGTAGCTATTTTTTTTAA
TCAATGTGTTGCTAGTTGCTGTTTATGGTGCTTTGACAGTTGATAAAAAAAA
TCTTATTAACAGGGGTGGTTTACCTATATTAGCTCTTTTAACATTCTTATT
35 TTAA

(SEQ ID NO: 242)

TFLYYVFEDVATQSNMTGKIFSMSKEELSYLPVIKLFKNQGVYNGLIGLFLLY
GLYISQNQEIVAIFLINVLLVAVYGALTVDKKILLKQGGLPILALLTFLF*

40

Sequence description:

A] Length: 312 bp - 104 aa (partial gene sequence)

B] This gene sequence was not identified using the LEEP system. It was identified upstream of the ID-93 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-93 gene sequence.

5 N-terminus has yet to be determined

ID-199

10 Clone RS-70c (ID-93c)

(SEQ ID NO: 243)
ATGAAATTAAGTGTCTTGATTATGGGCTTATTGATTATGGAAAACTGCA
AGTGATGCAATACAAGAAACGATTCTTTTATCACAAGAGGCGGAGCAACT
AGGCTATCATCAATTTTGGGTGGCTGAACATCACGGTGTTAAGGCATTTCAG
15 TATTAGCAATCCAGAATTAATGATAATGCATTTGGCTAACCAGACTAAATC
TATCAAAATTGGCTCTGGAGGTATAATGCCTCTGCACTATAGTAGTTTTAA
ACTCGCGGAGACTCTCAAGACATTAGAGACATGTCATCCGAATCGAGTAA
GTATTGGTTTAGGAAATTCAGTAAAGTTTCAAATGCACTTC
GTAGCTTACATAAAGCACATGATTACGAAGAGGTACTGGAGGAATTGAAG
20 TCATGGCTTATTGATGAATCATCCAGTAAGGAACCATTAGTTCAACCGACT
CTTTCTAGCTTCCCAGACTTATATGTGTTGGGGAGTGGTCAAAAATCAGCT
TATTTAGCGGCTAAACTTGGCTTAGGCTTTACCTTCGGTGTTTTTCCTTTTA
TGGACAAAGACCCATTGACAGAAGCTAAA

(SEQ ID NO: 244)
25 MKLSVLDYGLIDYGKTASDAIQETILLSQEAEQLGYHQFWVAEHHGVKAFFSIS
NPELMIMHLANQTKSIKIGSGGIMPLHYSSFKLAETLKTLETCHPNRVSIGLGN
SLGTVKVSNALRSLHKAHDYEEVLEELKSWLIDESSSKEPLVQPTLSSFPDLYV
LGSGQKSAYLAACLGLGFTFGVFPFMDKDPLTEAK

30

Sequence description:

A] Length: 588 bp - 196 aa (partial)

35 B] This gene sequence was not identified using the LEEP system. It was identified downstream of the ID-93 gene which was identified by LEEP, during cloning and sequence analysis of the full-length ID-93 gene sequence. No obvious signal peptide, but Shine Dalgarno sequence upstream of the ATG start codon.

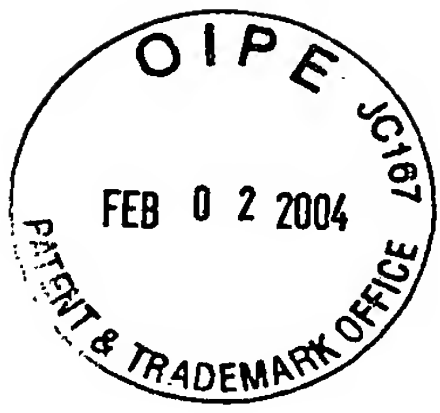


Figure 3

nucS1

Bgl II Eco RV

5'-cgagatctgatatctcacaaacagataacggcgtaaataag -3' (SEQ ID NO: 245)

nucS2

Bgl II Sna I

5'-gaagatcttcccccggagatcacaacagataacggcgtaaataag -3' (SEQ ID NO: 246)

nucS3

Bgl II Eco RV

5'-cgagatctgatatccatcacaacagataacggcgtaaataag -3' (SEQ ID NO: 247)

nucR

Bam HI

5'-cgggatcccttatggacctgaatcagcgttgtc -3' (SEQ ID NO: 248)

NucSeq

5'-ggatgctttgtttcaggtgtatc -3' (SEQ ID NO: 249)

PTREP_F

5'-catgatatcggtacctcaagctcatatcattgtccggcaatgggtgtgggctttttttgttttagcggataa
caatttcacac -3' (SEQ ID NO: 250)

PTREP_R

5'-gaggatcccccgggcttaattaatgtttaaacactagtcgaagatctcgcgaattctcctgtgtgaaatt
gttatccgcta -3' (SEQ ID NO: 251)

PUC_F

5'-cgccagggttttccccagtcacgac -3' (SEQ ID NO: 252)

V_R

5'-tcagggggggcggagcctatg -3' (SEQ ID NO: 253)

V₁

5'-tcgtatgtttgtgtggaattgtg -3' (SEQ ID NO: 254)

V₂

5'-tccggctcgtatgtttgtgtggaattg -3' (SEQ ID NO: 255)

Figure 4

pTREP-Nuc vectors allow cloning of genomic DNA into each frame with respect to the nuclease gene

5

(i)

pTREP1-nuc1 (EcoRV)	AAGTATCAGATCT--GATATC--TCACAAACAGATAACGGCGTAAAT	Frame=+1
(SEQ ID NO: 256)▲.....	
pTREP1-nuc2 (Sma I)	AAGTATCAGATCTTCCTCCGGA-TCACAAACAGATAACGGCGTAAAT	Frame=+2
(SEQ ID NO: 257)▲.....	
pTREP1-nuc3 (EcoRV)	AAGTATCAGATCT--GATATCCATCACAACAGATAACGGCGTAAAT	Frame=+3
(SEQ ID NO: 258)▲.....	
Nuclease Gene	TCACAAACAGATAACGGCGTAAAT	
	(SEQ ID NO: 259)	
Cloning site is indicated by an arrow		

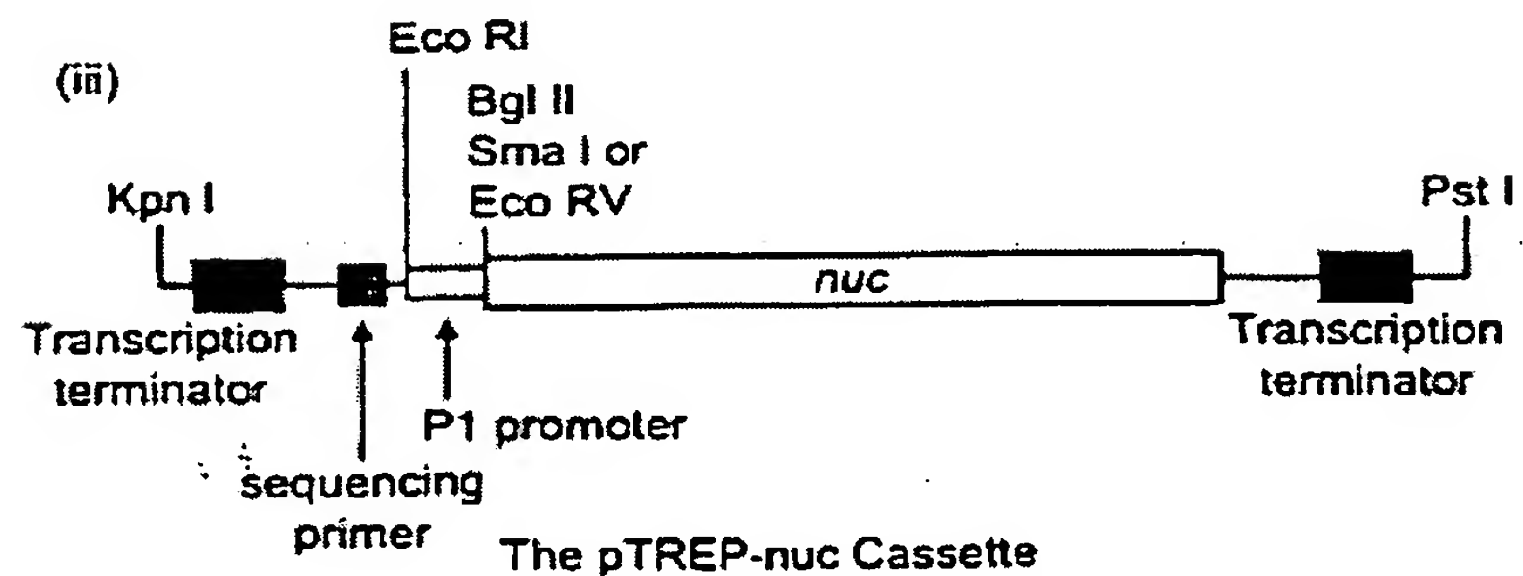
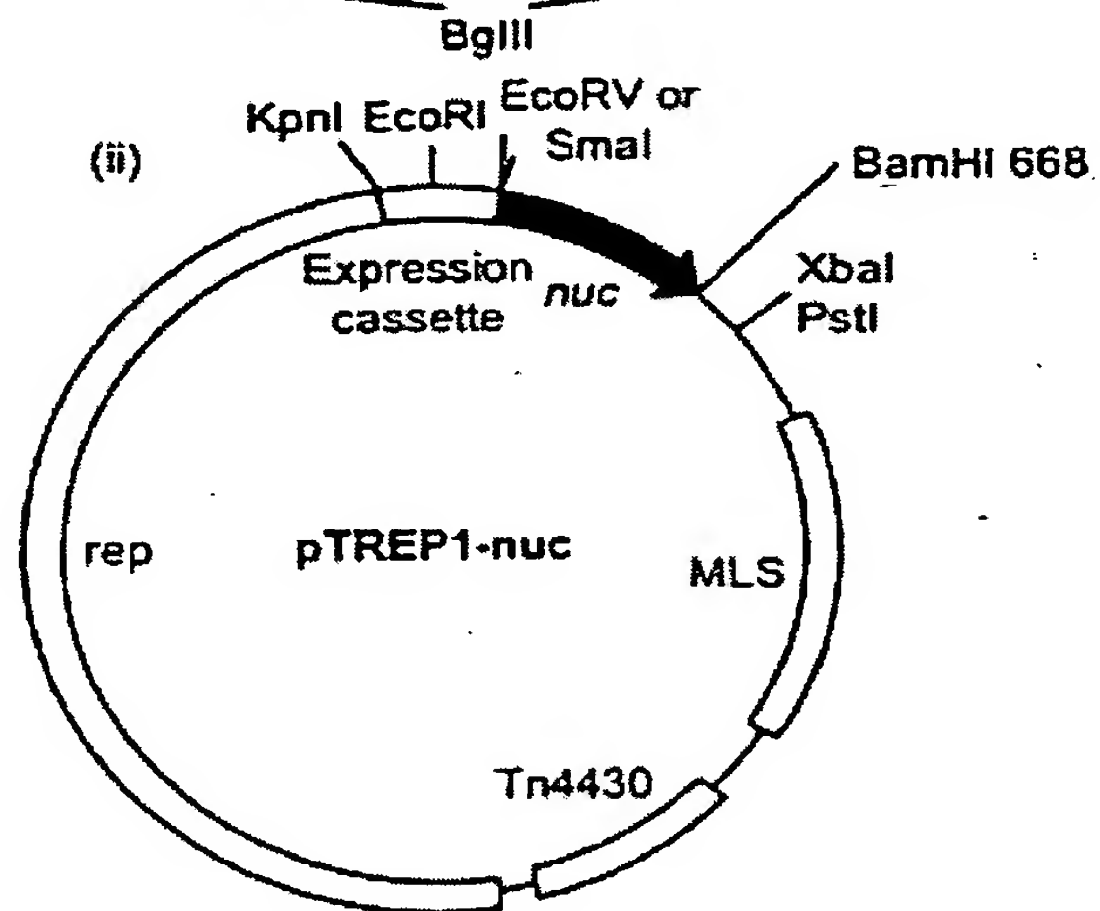
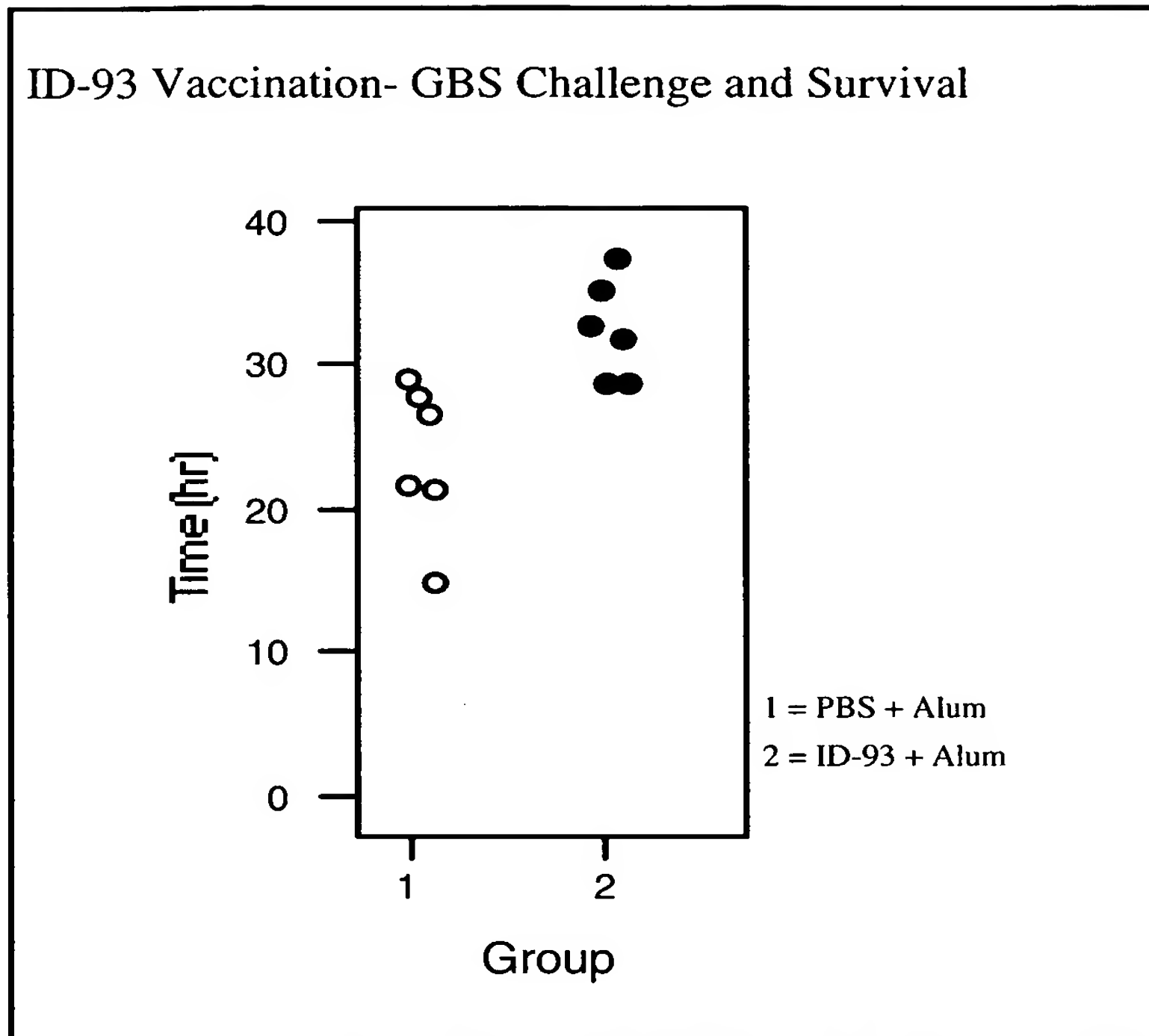


Figure 9
Survival Data



~~FIGURE 9 SURVIVAL DATA~~